Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Department of Psychology, University of Michigan
Ann Arbor, Michigan

DAVID A. GRANT, Associate Editor
University of Wisconsin

DELOS D. WICKENS, Associate Editor
Ohio State University

Consulting Editors

E. JAMES ARCHER, *University of Wisconsin*
JUDSON S. BROWN, *State University of Iowa*
CLETUS J. BURKE, *Indiana University*
PAUL M. FITTS, *Ohio State University*
FREDERICK C. FRICK, *Massachusetts Institute of Technology*
FRANK A. GELDARD, *University of Virginia*
JAMES J. GIBSON, *Cornell University*
CLARENCE H. GRAHAM, *Columbia University*
HAROLD W. HAKE, *University of Illinois*
LLOYD G. HUMPHREYS, *AF Personnel and Training Research Center*
ARTHUR L. IRION, *Tulane University*
HOWARD H. KENDLER, *New York University*
DONALD B. LINDSLEY, *University of California, Los Angeles*
KENNETH MACCORQUODALE, *University of Minnesota*
QUINN MCNEMAR, *Stanford University*
NEAL E. MILLER, *Yale University*
KENNETH W. SPENCE, *State University of Iowa*

ARTHUR C. HOFFMAN, Managing Editor
HELEN ORR, Assistant Managing Editor
Editorial Staff: SADIE J. DOYLE, FRANCES H. CLARK, BARBARA CUMMINGS

CONTENTS

Transfer of Experience With a Class-Schema to Identification-Learning of Patterns and Shapes: F. ATTNEAVE 81

Extinction as a Function of Frustration Drive and Frustration Drive Stimulus: B. B. BERNSTEIN 89

Some Factors Affecting Performance on a Word-Formation Problem: W. F. BATTIG 96

Effect of Addition of Irrelevant Verbal Cues on Perceptual-Motor Learning: H. W. BRAUN AND A. W. BENDIG 105

Operator Performance as a Function of the Statistical Encoding of Stimuli: W. D. GARVEY 109

Expectation and Resistance to Extinction of a Lever-Pulling Response as Functions of Percentage of Reinforcement and Amount of Reward: D. J. LEWIS AND C. P. DUNCAN 115

Effects of Counting and Ordering Habits on the Acquisition of a Simple Motor Skill: J. H. BOWEN, T. G. ANDREWS, AND S. ROSS 121

Continuous Perspective Transformations and the Perception of Rigid Motion: J. J. GIBSON AND E. J. GIBSON 129

Studies of Distributed Practice: XVI. Some Evidence on the Nature of the Inhibition Involved in Massed Learning of Verbal Materials: B. J. UNDERWOOD 139

Two Determinants of Associative Reaction Time: J. S. WIGGINS 144

The Effect of an Irrelevant Drive on Maze Learning in the Rat: H. W. BRAUN, C. E. WEDEKIND, AND J. F. SMUDSKI 148

American Psychological Association

Vol. 54 No. 2 August 1957

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $8.00, or $16.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal, experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research effort is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this JOURNAL, 1957, 53, 1–2).

Address all articles submitted for publication to the editor, Arthur W. Melton, Department of Psychology, University of Michigan, Ann Arbor, Michigan. All manuscripts should be submitted in two complete copies; only the original copy will be returned if the article is not accepted for publication. Figures may be original line drawings or glossy photographic prints, but should not exceed 8½ by 11 in. Care should be taken to insure legibility of labels and critical features of drawings when reduced to a single-column or (exceptionally) a double-column figure in the JOURNAL.

To save printing costs supplementary material to articles in this JOURNAL are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 [4, Suppl.], 389–449). In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Ss*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879. Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1957 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 54, NO. 2 AUGUST, 1957

TRANSFER OF EXPERIENCE WITH A CLASS-SCHEMA TO IDENTIFICATION-LEARNING OF PATTERNS AND SHAPES [1]

FRED ATTNEAVE

Operator Laboratory, Air Force Personnel and Training Research Center

The idea that experience provides the organism with some "apperceptive mass" which facilitates later perception and learning is an old one, with many variants and elaborations. The concept of a "schema" as an entity mediating the effects of past experience has been prominent in recent British psychology, chiefly because of the thinking of Bartlett (**5**) and his associates, as reflected in the summary and review of Oldfield and Zangwill (**12**). Woodworth (**13**) and Hebb (**7**) have used the term "schema" in a sense which is somewhat more restricted and more definite than Bartlett's. After considering a number of experiments on memory for form, Woodworth concluded that a new configuration is usually remembered in terms of a "schema, with correction." For example, a figure which may be described as "a square with a nick on one side" is easier to learn than most other seven-sided polygons because the schema "square" is simple, familiar, and unambiguous, and the correction "with a nick in one side" is easily and clearly specifiable. Hebb emphasizes the importance of acquiring the schema of a class of objects to be differentiated from one another: thus Chinese "all look alike" to the occidental observer who has not seen many Chinese. So conceived, the schema consists, at least in part, of some representation of the central tendency or communality of the class of objects in question. If the observer has some subjective standard of the human face which he has obtained by "averaging" the faces of Americans, he may learn a new American face in terms of the manner and degree in which it deviates from this schema (cf. Woodworth's "correction"). If he is suddenly thrust into a Chinese population, however, his standard will no longer be central, and the new faces will all deviate from it in more or less the same direction.

Recently Oldfield (**11**), Hochberg (**8**), and the writer (**3**) have separately pointed out that the use of schemata makes for economical information storage. Oldfield's development of

[1] This report is based on work done under ARDC Project No. 7706, Task No. 27001, in support of the research and development program of the Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government.

this idea is the most detailed: Whatever suspicion one might have that the schema concept need be subjective and amorphous is removed by his demonstration that a computer might employ schemata (in essentially the Woodworth-Hebb sense) to good advantage.

In view of the theoretical attractiveness of this concept, it is surprising that little or no directly relevant experimentation has been undertaken. To the best of the writer's knowledge, the two experiments reported here constitute the first test under controlled conditions of the hypothesis, most directly attributable to Hebb, that the learning of a class schema makes easier the subsequent learning of identifying responses to members of the class.

Experiment I: Letter-Patterns

Method

Materials.—The experimental stimuli consisted of three prototype patterns, and eight variations on each prototype. A *prototype* was constructed in the following manner. In a matrix six units wide and five high, each cell was either filled with a letter of the alphabet or left blank. For each cell, the probability of a blank was ½, and the probability of any particular letter was 1/52. The state of each cell was independently determined by the use of a table of random numbers, in accordance with these probabilities. A *variation* was a pattern in which one letter of the prototype was changed to some other letter (blanks were never varied). Both the position of the letter changed and the identity of the letter to which it was changed were randomly chosen. Figure 1 shows one of the three prototypes and four of the eight random variations on it. Note that the variations necessarily differ from one another more than they differ from the prototype.

Fig. 1. Letter patterns used in Exp. I. A prototype (bottom) and 4 of the 8 variations on it.

In the actual stimulus patterns, lines marking off rows and columns were omitted, as in Fig. 1. The patterns were typed with an electric typewriter (single spacing between rows; double spacing between columns) on a fine-grain, translucent tracing paper, behind which a reversed sheet of carbon paper was placed to increase the density of the typing. The patterns were then cut out and mounted in 2 × 2 in. slides, with due attention to uniformity of framing. Although the paper somewhat reduced the brightness of the ground, the projected images were quite satisfactory with respect to contrast and legibility.

Subjects.—A total of 60 airmen (basic trainees) at Lackland Air Force Base served as *Ss*.

Procedure.—The 30 *Ss* assigned to Group E were given pretraining which consisted of repeated viewings of one of the prototype patterns, interspersed with attempts to reproduce it, whereas the 30 *Ss* of Group C were given an equivalent amount of practice in the reproduction of a completely irrelevant figure (a nonsense shape). Both groups were then tested on a paired-associates learning task in which the stimulus members were variations on the prototype used in the pretraining of Group E.

Pretraining consisted of eight reproduction trials. The pretraining stimulus was exposed (projected to the wall in front of *S*) for 15 sec. each time; after each exposure *S* attempted to put on paper as accurate a reproduction of the stimulus as he could. In the paired-associates learning task, the stimuli were eight variations on the prototype reproduced by Group E, and the responses were eight names of men (three letters each; no two with the same initial letter; e.g., *Sam*, *Joe*). Six learning trials were given, each followed by a test trial. Twelve different random permutations were used for these 12 presentations of the stimuli. All stimulus exposures were of 10 sec. duration, with 5 sec. between exposures on the same trial: the slide projector was automatically stepped at this rate by a set of Hunter timers. On learning trials, the "name" of each pattern was given aloud from a tape recording as the pattern was projected (the tape and the timing mechanism of the projector were synchronized by *E* at the beginning of the

trial). On test tri[illegible] attempted to write the correct name of eac[illegible]ttern, as it appeared, in a mimeographed boo[illegible] which provided an individual sheet for every trial. The interval between successive trials was 1 min., during which *E* rearranged the slides for the next presentation.

The *Ss* were taken in subgroups of 10 at a time. Each of the three subgroups making up Group E was given a different prototype and associated variations. The same three sets of variations were likewise assigned to the three subgroups of Group C.

At no time was *S* told that the pretraining activity might improve his performance on a subsequent task; instead he was given the impression that the reproduction and the identification learning were two separate "memory tests." The only mention of a connection between the two was made to Group E, near the beginning of the instructions preceding paired-associates learning: "Now I'm going to give you a different kind of memory test. I'm going to show you eight patterns, made up of letters of the alphabet. They are all pretty much alike (in fact, they are all variations of the pattern you've been copying) but they are different enough so that you can learn to tell them apart . . . ." The parenthetical clause was of course omitted for Group C.

The objective of providing experimental *Ss* with a subjective standard or schema corresponding to the prototype pattern might have been accomplished by various pretraining techniques other than reproduction. Arnoult (1) has found, however, that reproduction is the most effective of a wide variety of methods for learning a pattern, when the criterion of learning is the ability to recognize the pattern in a context of others very similar. Reproductions of the prototype by Group E were not systematically scored, but were inspected fairly thoroughly. Some *Ss* achieved formally perfect reproductions; typically, however, a few errors were still present on the last trial.

Results

Errors made during paired-associates learning were summated over all six trials for each *S*; these total scores were then averaged within groups. The mean error score was 36.20 for Group C and 28.47—about 21% less—for Group E. It should be added, however, that the task was altogether too difficult for a good many of the *Ss*: on the sixth and final trial, 18 *Ss* (5 in Group E, 13 in Group C) were still making purely chance scores of 7 or 8 errors. On the same trial, 13 experimental *Ss* and 6 control *Ss* made no more than 2 errors; the modal numbers of errors was zero (with 9 *Ss*) for Group E and 7 (with 8 *Ss*) for Group C.

Since the results showed no evidence of any real differences between subgroups given different specific patterns, subgroups were simply pooled in a t test of the difference between total errors of Group E and Group C. A value of $t = 2.76$ was obtained; with $df = 58$, this is significant at the .01 level.

Experiment II: Polygons

In the case of the second experiment, which employed angular nonsense shapes instead of letter patterns, the advantages of pretraining with a prototype were somewhat less obvious, by intuitive standards, than formerly. Also the design was more complex, introducing an additional variable having to do with the way in which the variations differed from the prototype and from one another.

Method

Materials.—Ten prototype shapes were constructed by plotting points with random coordinates in a 16×16 matrix, and then connecting the points with the shortest possible closed contour (4). Five 6-sided and five 12-sided polygons were thus prepared: all 10 are reproduced in the two upper rows of Fig. 2. In each of the variations on such a prototype, one-third of its points (2 out of 6, or 4 out of 12) were moved. The movement of a point was from its original matrix cell to a randomly chosen one of the eight immediately surrounding cells: i.e., the direction of movement varied randomly, but the extent of movement was approximately constant (either 1 or $\sqrt{2}$ matrix units). From each prototype, two sets of eight variations were made. In the first set, the *same* two or four points were varied throughout the whole set (though the selection of these points was initially random). In the

FIG. 2. The 10 shapes used as prototypes in Exp. II are shown in the two upper rows; a sample set of variations is shown in the two lower rows.

second, an independent random selection of the two or four points to be moved was made for each individual variation. In other words, the *loci* of variation were constant for the first set and randomly variable for the second.

The two lower rows of Fig. 2 show a set of variations on one of the prototypes. It follows from the method of construction that the variations "vary about" the prototype: i.e., that the prototype represents the central tendency of the class.

Each of the 170 shapes (10 prototypes and 20 sets of 8 variations) was mounted in a 2 × 2 in. slide. The figure-ground relationship was the reverse of that shown in Fig. 2, i.e., the shape was transparent and the remainder of the slide opaque. No matrix lines appeared on the slide.

Subjects.—The *S*s were 320 airmen from the same population as in Exp. I.

Procedure.—In all essential respects, the procedure was the same as in Exp. I. Subgroups were made up of 8 instead of 10 *S*s, an experimental and a control subgroup being employed with each of the 20 sets of variations. Experimental *S*s were again given pretraining which consisted of eight reproduction trials with the relevant prototype. Control *S*s were given pretraining with an unrelated shape: if their paired-associates stimuli were to be 6-sided polygons, they were pretrained with one of the 12-sided polygons, and vice versa. The first eight letters of the alphabet, instead of men's names, were used as response members of the paired associates. Number and duration of trials were as before.

Results

Differences in errors on the paired-associates task appeared on first analysis to be small and variable: 7 of the 20 matched pairs of subgroups showed differences in the unexpected direction (see Table 1). With the view of compensating for the high degree of inter-*S* variability, a predictor variable which might be used for covariance control was sought. The

TABLE 1

SCORES ON EXPERIMENTAL TASK AND ON TEST OF INITIAL ABILITY

Pairs of Subgroups	Mean Error		Test Score		r (Error-ML)	Unadjusted Error Diff.	Adjusted Error Diff.
	Exp.	Control	Exp.	Control			
6Sa	18.13	24.25	4.38	4.63	−.61	6.12	6.56
6Ra	28.25	29.75	3.13	4.38	−.72	1.50	6.42
6Sb	33.63	31.75	3.25	5.38	−.60	−1.87	3.44
6Rb	26.25	26.88	4.75	4.88	−.53	.62	.86
6Sc	35.88	34.00	3.75	4.25	−.42	−1.87	−1.16
6Rc	29.63	27.00	3.88	4.63	−.33	−2.62	−1.42
6Sd	26.38	29.88	3.88	4.38	−.73	3.50	5.09
6Rd	31.38	32.75	4.88	4.38	−.41	1.37	.58
6Se	24.38	27.75	4.88	4.38	−.58	3.37	2.31
6Re	27.63	26.25	4.63	3.88	+.01	−1.37	−1.34
12Sa	25.00	22.50	3.50	5.50	−.45	−2.50	3.56
12Ra	27.13	28.62	3.13	3.25	−.56	1.50	1.94
12Sb	28.00	30.38	4.75	5.00	−.56	2.37	3.48
12Rb	29.25	23.13	4.50	6.75	−.39	−6.12	−2.62
12Sc	27.50	30.50	3.88	4.75	−.83	3.00	6.53
12Rc	20.50	30.63	5.63	4.63	−.41	10.12	7.41
12Sd	28.13	29.25	2.75	4.13	−.47	1.12	4.54
12Rd	23.00	27.00	5.75	5.13	−.17	4.00	3.49
12Se	27.13	28.13	3.88	5.00	−.32	1.00	2.84
12Re	26.13	20.88	3.25	5.00	−.54	−5.25	.97
Mean	27.17	28.06	4.12	4.72	−0.52*	.90	2.67

* Root-mean square.

classification battery given to all Air Force basic trainees contains a test known as "Memory for Landmarks" which is not unlike the identification-learning task of the present experiment, and which showed a correlation with the present task of $r = -.52$ (RMS average of within-subgroups correlations). Moreover, control *Ss* were found to have been disconcertingly higher in initial ability, by the "Memory for Landmarks" criterion, than experimental *Ss* (see Table 1); thus, some control for initial ability was even more strongly indicated.

Now, the experiment was originally so designed that an error term might be based upon variability among subgroups differing in both *Ss* and stimuli, in order that "significant" results might be considered to hold for a parent population of stimuli, as well as for a parent population of *Ss* (**4**). By adhering to this strict criterion of significance, and excluding from consideration any weaker types of generalization, it was possible to handle the covariance adjustment and the subsequent tests in an unusually simple and economical manner.

The data were first simplified by obtaining difference scores for matched *E* and *C* subgroups (see Table 1). An adjustment for initial ability was then made independently on each of these 20 difference scores, by the formula

$$\text{Adjusted } D_x = D_x - b_{xy} D_y,$$

in which D_x is the difference between the subgroup means on the experimental task, D_y is the difference on the "Memory for Landmarks" test, and b_{xy} is the within-subgroups regression coefficient. It should be emphasized that b_{xy} was obtained separately for

TABLE 2

ERROR-DIFFERENCES BETWEEN COVARIANCE-ADJUSTED MEANS OF MATCHED EXPERIMENTAL AND CONTROL SUBGROUPS

Sides	Prototype	Points Varied		S+R	S−R
		S(ame)	R(andom)		
6	6a	6.56	6.42	12.98	.14
	6b	3.44	.86	4.30	2.58
	6c	−1.16	−1.42	−2.58	.26
	6d	5.09	.58	5.67	4.51
	6e	2.31	−1.34	.97	3.65
12	12a	3.56	1.94	5.50	1.62
	12b	3.48	−2.62	.86	6.10
	12c	6.53	7.41	13.94	−.88
	12d	4.54	3.49	8.03	1.05
	12e	2.84	.97	3.81	1.87
Mean		3.72	1.63	5.35	2.09
t ($df = 9$)		5.25	1.56	3.25	3.06
P (2-tail test)		<.001	<.20	=.01	<.02

Note:—In Columns S, R, and S+R, a negative sign indicates that control *Ss* made fewer errors than experimental *Ss*. Each t is a test of the null hypothesis that the column mean is a chance deviation from zero.

each pair of subgroups to allow for the possibility of true differences in correlation dependent on stimulus materials; also because it was considered desirable that the adjusted difference scores be completely independent of one another with respect to statistical restraints. Table 1 shows correlation coefficients rather than regression coefficients; the relation between the two is: $b_{xy} = r\frac{\sigma_x}{\sigma_y}$ (**10**, Ch. 15).

The adjusted difference scores are classified by systematic stimulus characteristics in Table 2. These 20 scores are still not entirely independent with respect to stimuli, because two sets of variations (one with the same points moved, the other with random points moved) were derived from each of the 10 prototypes. A set of 10 completely independent scores (the "S + R" column of Table 2) was obtained by combining difference scores associated with the same prototype. A t test of the departure of this distribution from zero yields a value of 3.25, with $df = 9$, for which P is just equal to .01. This measures our confidence that the over-all difference between Group E and Group C is generalizable both to other *Ss* and to other stimuli constructed by the same rules.

Considering results on "Same" and "Random" variations separately, in the former case a t of 5.25 is found with $df = 9$, significant at the .001 level, and in the latter case there is a nonsignificant t of 1.56 in the expected direction. The "S − R" column of Table 1 contains the distribution of differences between difference scores obtained with "Same" and "Random" variations, matched by prototype. The t for the departure of this distribution from zero is 3.06 with $df = 9$, significant at the .02 level. One may be reasonably sure, therefore, that pretraining with the prototype did more good when the same parts of the figure were subsequently varied than when different parts were varied. A plausible interpretation of this result, in crude terms, is as follows: When an *S* who has become familiar with the prototype sees the first of the variations, his attention is drawn to those parts of the figure which differ from the prototype, and a set with respect to *where to look* for distinguishing characteristics is thereby established. This set is appropriate and useful if the loci of variation are constant throughout the class, but not otherwise. One would be ill-advised, however, actually to accept the null hypothesis associated with the nonsignificant t for the "Random" condition, and to conclude that familiarity with a prototype is effective *only* when the loci of variation are constant, for at least two reasons: First, and most im-

portant, positive results were obtained in Exp. I, in which locus of variation was not constant. Second, the actual odds against obtaining by chance a t as great as the one in question (t = 1.56, df = 9) are about 6 to 1, and conventions having to do with "significance" need not prevent us from giving these odds their due weight.

The experimental effect did not appear to vary with stimulus complexity: a comparison of "S + R" values for 6- vs. 12-sided polygons yielded a t of only .37 with df = 8. A similar test on the "S − R" column (essentially a test for interaction) likewise yielded a nonsignificant t = .06.

Discussion

Both of the experiments reported involve a possible attenuation of effects which should be considered carefully in the evaluation of the results. It may not be assumed that control *Ss* learned to identify the variations without the benefit of a class schema, since there was nothing to prevent them from abstracting the schema once they were exposed to the class. Indeed, the very principle on which one would predict superiority of Group E over Group C would normally apply to situations, not unlike that confronted by the control *Ss*, in which schemata are available only by abstraction from natural objects. The most one can assume is that the pretraining of experimental *Ss* gave them a head start over the control *Ss* in the acquisition of an appropriate schema. Accordingly, the obtained differences in favor of Group E can be considered no more than dilute manifestations of the effect of schemata on identification learning.

It is evident that a close relationship must exist between the mechanisms associated with schemata and with "stimulus-predifferentiation," as studied by the Gibsons and others (2). The present experiments do not involve predifferentiation in any literal sense, for the pretraining was on a single stimulus; on the other hand one may justifiably assume that predifferentiation training does involve schema learning. This is not to say that the predifferentiation subject acquires nothing more than a knowledge of the "average" stimulus. More likely, he learns something about at least three characteristics of the class: (*a*) its central tendency; (*b*) *how* its members may differ from one another, i.e., in what properties, or on what dimensions; and (*c*) its dispersion, i.e., *how much* its members may differ from one another on the several dimensions of variability. The importance of (*a*) is of course indicated in the present studies. Sensitivity to both (*a*) and (*c*), in unidimensional situations, is implied by the relativity of judgment which individuals commonly display in assigning stimuli to rating-scale categories. The importance which (*b*) assumes in multivariate situations has recently been shown by Kurtz (9) (cf. also the difference between the "Same" and "Random" conditions in Exp. II). Although in the present paper *schema* has been used in the Woodworth-Hebb sense, to refer to some representation of the central tendency of a multivariate class, there is perhaps no reason why its reference should not be extended to include the whole system of class parameters suggested above.

The importance of schemata in learning is supported by so much commonsense evidence that the present studies may well be considered demonstrations rather than experiments. It is hoped, however, that these studies may open the way for others investigating more uncertain matters. A particularly interesting question concerns the effect of pretraining with a noncentral standard, i.e., one which is similar but peripheral to the stimuli of the paired-associates class. Positive transfer might be predicted on the ground that the standard shares information with the paired-associates stimuli; alternatively, negative transfer might be predicted on the ground that the stimuli to be identified all deviate in the same direction from the standard, and therefore should appear more alike.

The answer to this question would throw some light on the mechanisms underlying the present findings.

Summary

Two experiments were conducted to determine the effect on a paired-associates learning task of prior familiarization with a single "prototype" stimulus representing the central tendency of the stimuli to be identified. The stimuli were letter patterns in Exp. I, and polygons in Exp. II. In both cases they were constructed by systematically random methods.

In Exp. II a sufficient sample of shapes was used to permit statistical generalization of the results to other stimuli as well as to other *Ss*. Two different methods were employed in this experiment to generate the paired-associates stimuli: in one, the same parts of the shapes were varied throughout the set; in the other, different parts were varied from shape to shape.

Positive results were obtained in both experiments: i.e., familiarization with the central "prototype" stimulus decreased errors on the paired-associates task. In Exp. II this effect was greater when the loci of variation among the shapes were constant.

References

1. Arnoult, M. D. A comparison of training methods in the recognition of spatial patterns. Lackland AFB, Tex.: Air Force Personnel and Training Research Center, February, 1956. (*Res. Rep.* AFPTRC-TN-56-27.)
2. Arnoult, M. D. Stimulus predifferentiation: some generalizations and hypotheses. *Psychol. Bull.*, 1957, **54**, 340–351.
3. Attneave, F. Some informational aspects of visual perception. *Psychol. Rev.*, 1954, **61**, 183–193.
4. Attneave, F., & Arnoult, M. D. The quantitative study of shape and pattern perception. *Psychol. Bull.*, 1956, **53**, 452–471.
5. Bartlett, F. C. *Remembering, a study in experimental and social psychology.* Cambridge: Cambridge Univer. Press, 1932.
6. Guilford, J. P. (Ed.) *Printed classification tests.* Washington: U. S. Govt. Print. Off., 1947. Army Air Forces Aviation Psychology Program Res. Rep. No. 5.
7. Hebb, D. O. *Organization of behavior.* New York: Wiley, 1949.
8. Hochberg, J. E. Psychophysics and sterotypy in social perception. In *Emergent problems in social psychology* (in press).
9. Kurtz, K. H. Discrimination of complex stimuli: the relationship of training and test stimuli in transfer of discrimination. *J. exp. Psychol.*, 1955, **50**, 283–292.
10. McNemar, Q. *Psychological statistics.* New York: Wiley, 1949.
11. Oldfield, R. C. Memory mechanisms and the theory of schemata. *Brit. J. Psychol.*, 1954, **45**, 14–23.
12. Oldfield, R. C., & Zangwill, O. L. Head's concept of the schema and its application in contemporary British psychology. *Brit. J. Psychol.*, 1942, **32**, 267–286; 1943, **33**, 58–64, 113–129, and 143–149.
13. Woodworth, R. S. *Experimental psychology.* New York: Holt, 1938.

(Received July 3, 1956)

Journal of Experimental Psychology
Vol. 54, No. 2, 1957

EXTINCTION AS A FUNCTION OF FRUSTRATION DRIVE AND FRUSTRATION-DRIVE STIMULUS [1]

BENJAMIN B. BERNSTEIN [2]

University of Missouri

The relationship between the blocking of a response and the subsequent increase in vigor of response has been interpreted by many investigators to be a function of frustration drive (**1**, **2**, **4**, **7**, **8**). Applying current drive theory, a number of these authors have posited a frustration stimulus (S_F), which would accompany the drive state and be capable of mediating the transfer of responses or response tendencies from one stimulus situation to another. This study investigated both components of frustration. Two experimental situations were used: (*a*) a frustration-drive test was designed to determine the effect of blocking an avoidance response on response vigor and resistance to extinction; and (*b*) a frustration-stimulus test was designed to test for the occurrence of response transfer mediated by S_F.

Frustration-drive test.—Although there is evidence that the occurrence of response blocking will lead to an increase in response vigor, no systematic attempts have been made to determine the effects of blocking an avoidance response during extinction. This study was designed to examine some of these effects, and to compare them with response predictions based on the frustration-drive hypothesis.

In the frustration-drive test, an avoidance response (the turning of a wheel that eliminated shock) was blocked for different lengths of time during extinction. It was predicted from frustration-drive theory that an increase in response vigor and resistance to extinction would occur for short durations of blocking time. However, because of the assumed dependence of frustration on initial level of drive intensity (**2**), frustration effects were expected to diminish when response blocking exceeded a critical duration.

Frustration-stimulus test.—On the day following extinction of the shock-avoidance response, the same *S*s learned a completely different response: running to food in a straight alley. The resistance to extinction of this response was used to test for the transfer of response tendencies learned as a result of blocking on the previous day.

The prediction of response transfer was based on the hypothesis that S_F accompanies the occurrence of frustration drive. Because an extinction trial in the runway may be considered to be one in which the goal response (R_G) is blocked, S_F's were expected to occur and thus provide mediation for the transfer of responses associated with S_F in avoidance extinction. It was reasoned that if persistent goal-directed activity in response to a frustrating condition is rewarded in one situation, response persistence would tend to occur in other frustrating situations, regardless of external stimulus differences. The demonstration of this would provide experimental support for the use of frustration drive stimulus as an explanatory concept.

METHOD

Subjects.—Forty-four male and 60 female albino rats between 90 and 120 days of age were used. All were drawn from the colony main-

[1] This report is based on a dissertation submitted to the Graduate School of the University of Missouri in partial fulfillment of the requirements for the Ph.D. degree. The author is indebted to Dr. Melvin H. Marx for his advice and criticism.

[2] Now with the Bombing-Navigation Research Unit, Operator Laboratory Field Unit No. 4, Air Force Personnel and Training Research Center, Mather Air Force Base, California.

tained by the Agricultural Chemistry Department of the University of Missouri. In the frustration-drive test, eight *Ss* were discarded for failing to reach criterion; one *S* was discarded because of a procedural error; and one *S* was not used in data analysis because it was a statistical deviate.[3] For failing to reach criterion in the frustration-stimulus test, three *Ss* were lost from one group, and four *Ss* were lost from each of the remaining four groups.

Apparatus.—A modified Mowrer-Miller box (5) was used in the frustration-drive test, and a straight-alley runway was used in the frustration-stimulus test.

Half of the Mowrer-Miller box was sectioned off by a partition of unpainted plywood. A small wheel (4-in. diameter) available to *Ss* was the only manipulandum present. A sheet-metal guillotine door was constructed so that it could be operated to block the avoidance response (wheel-turning). The intensity of shock was 150-v. A.C., and a 410,000-ohm resistance was in series with the rat. Two Hunter decade timers were used to time the interstimulus interval. The glass front of the Mowrer-Miller box faced the center of the experimental room, which was illuminated by a 100-w. ceiling lamp.

The straight-alley runway was 6 ft. long, 3 in. wide, and 5 in. deep. Two guillotine doors separated the runway from the starting and goal boxes. These boxes were each 12 in. long, 3 in. wide, and 5 in. deep. A black curtain was located at the goal-box door. The top of the runway and boxes was covered by hardware cloth, and the wood was painted a flat black. Two 60-w. lamps, each located 2 ft. above a guillotine door, illuminated the apparatus.

Two Standard Electric timers, calibrated in .01 sec., were used to obtain response times in the straight-alley runway. A microswitch attached to the starting-box door automatically started both electric timers. When *S* left the starting box, *E* stopped one of the timers by means of a switch. The second timer was stopped in the same manner when *S* entered the goal box.

The experimental food was moist ground stock feed. It was placed in a glass cup 2 in. high and 2 in. in diameter. The side of the cup, located at the rear of the goal box, was covered by masking tape.

Individual wooden boxes (11 in. long, 7 in. wide, and 6 in. deep) were used for feeding. The boxes were made of unpainted plywood and were covered by hardware cloth. At the entrance to each box was a guillotine door. The feeding boxes were used to facilitate adaptation of *Ss* to the goal box in the runway.

[3] Using Thompson's test (11), the *t* of 3.38 is significant beyond the .01 level.

Maintenance and adaptation to apparatus.—In order to keep *Ss* from competing for the food and to adapt them to the guillotine door, they were placed on a 23½-hr. deprivation schedule for six days and were fed individually for 30 min. Just before feeding time on the sixth day, each *S* was placed in the straight runway and allowed to explore for 4 min. During exploration all doors were open and the curtain was not hung.

Eight hours after this feeding, each *S* was run in the frustration-drive test. A series of five nonreinforced presentations of the CS (buzzer) during a 4-min. adaptation period constituted pretraining in the Mowrer-Miller box.

Frustration-drive test.—Pretraining, training, and extinction were completed during the same session. Immediately following the adaptation period in the Mowrer-Miller box, each *S* was given a series of avoidance trials spaced approximately 60 sec. apart. On each trial the CS (buzzer) and the UCS (shock) were turned on in succession with a 3-sec. interstimulus interval. The *S* was able to turn off the shock by rotating the wheel. If the response was made during the 3-sec. interstimulus interval, the shock was avoided. The CS was turned off 3 sec. after the occurrence of the avoidance response. Acquisition trials were continued until *S* reached a criterion of 10 successive CR's with latencies less than 3 sec. Three *Ss* were run successively during the same experimental session.

The assignment of rats to extinction conditions was done at random before experimental runs. Because 60 *Ss* (15 *Ss* in each group) were needed for use in the second experimental situation, the number of *Ss* in the frustration-drive test depended upon *Ss'* satisfying the criterion for running in the frustration-stimulus test. The four groups in the frustration-drive test were as follows:

Group	*N*	*Condition*
0-sec.	19	"No" response blocking
2-sec.	18	Response blocked for 2 sec.
4-sec.	19	Response blocked for 4 sec.
8-sec.	19	Response blocked for 8 sec.

During extinction in the 0-sec. group, the guillotine door was lowered and raised within approximately 1 sec. Onset of the CS occurred immediately after the door was raised. The CS remained on for 10 sec. without shock on each trial.

During extinction in the 2-sec. group, the guillotine door was lowered approximately 1 sec. before the onset of the CS. The CS remained on for 2 sec. before the wheel was made available to *S* when *E* lifted the guillotine door. The CS continued for 10 sec. after the door was raised;

during this time, the CR was not blocked. Except for the change in response-blocking time, the same procedure was followed for the 4-sec. and 8-sec. groups. Thus, the length of time (10 sec.) during which the CS was coincident with the availability of the wheel was the same for all groups.

As in training, trials were spaced approximately 60 sec. apart. No trial was begun without a period of at least 3 sec. during which *S* was not turning the wheel. Extinction was continued to a criterion of five consecutive avoidance-response latencies that were greater than 10 sec.

Frustration-stimulus test.—All *S*s on Day 1 served as *S*s on the following day in the runway. Another group of 15 *S*s was added as a control group. The control *S*s were assigned at random before experimental runs in the same manner as *S*s in the other groups. Thus, there were five groups on Day 2, with 15 *S*s in each.

Twenty minutes following extinction trials on Day 1, *S* was fed for 10 min. in an individual feeding box; 20 hr. later *S* was run in the straight alley. Training and extinction were completed in the same session. The *S* was given 20 acquisition trials under 20-hr. deprivation with moist stock feed as reward. He was allowed to eat in the goal box for 15 sec. and after each trial was returned to an individual retention cage for an intertrial interval of from 4 to 6 min. The *S*s not entering the goal box in 5 min. during acquisition were discarded.

The *S* was placed in the starting box facing the starting-box door. After approximately 3 sec., the door of the starting box was raised. This was done only with *S* facing the door. When *S* left the starting box, the door was lowered to prevent retracing; when *S* entered the goal box, the goal-box door was lowered. The operation of the microswitch, attached to the starting-box door, and the electric timer produced a slight click when the starting-box door was raised. The hum of a large fan, which continued throughout the experiment, tended to mask such noises. Because these noises were constant among groups, it is believed that the interpretation of the results is not affected.

After the last training trial, *S* was given 20 extinction trials. During extinction, an empty food cup was present in the goal box; and *S* was left in the goal box for 15 sec. If *S* did not enter the goal box within 2 min., he was removed from the runway to await the next trial. The intertrial interval during extinction was approximately the same as that in acquisition.

The response time was measured from the opening of the starting-box door to *S*'s entrance into the goal box, exclusive of tail.

Results

Frustration-Drive Test

Acquisition.—The median latency of the last five trials of acquisition for each *S* was used as an estimate of response latency at the end of acquisition. The mean of these scores for all groups combined is 1.19; the median is 1.15. Using reciprocals of the scores, a test of significance among the four groups is not significant ($F < 1$).

Extinction.—To measure response vigor, the sum of the wheel turns for each *S* was obtained for the first five trials of extinction. These trials were chosen because the effect of frustration on response vigor was expected to be greatest in the early trials of extinction, and because some *S*s extinguished in five trials. Since the distribution of these scores was skewed, a transformation to $\log_{10}(1 + X)$ was used.

The means and *SD*'s of the transformed scores are presented in Table 1. Figure 1 shows the predicted increase in response vigor, and the reversal of this trend with prolonged response blocking. An analysis of variance was used to test the significance of the curve's departure from linear regression. As presented by

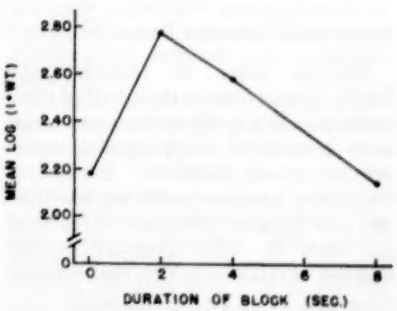

FIG. 1. Strength of wheel-turning (WT) response during avoidance extinction.

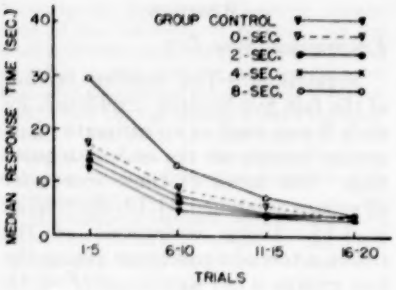

FIG. 2. Acquisition curves in runway.

Snedecor (**9**, pp. 379–381), the appropriate F is defined as the ratio of the mean square for curvilinear regression to the mean square for deviations from curvilinear regression. Using this test, a curvilinear relationship significant beyond the .05 level ($F = 4.61$, $df = 1$ and 72) was found.

The means and SD's of the number of trials to extinction for each group is shown in Table 1. A test of the linearity-of-regression hypothesis is significant beyond the .001 level ($F = 12.17$, $df = 1$ and 72).

No significant difference among the five groups (with control group included) was found in amount of food eaten during the 10 min. feeding following extinction trials in frustration-drive test.

Frustration-Stimulus Test

Because some *Ss* exceeded the 2-min. limit on more than half of their extinction trials, the median and mean were considered inappropriate measures of central tendency. Instead, a frequency measure, used by Sheffield (**8**) and Stanley (**10**), was calculated for each *S*. This measure is the number of trials on which the response time was at or below the median for all response times in all groups. This method was used to obtain an acquisition score and an extinction score for each *S*.

Acquisition.—The learning curves for the five groups in the runway are shown in Fig. 2. The median response times of successive blocks of five trials could be used to show learning because in acquisition no limit on response time was imposed.

To test acquisition differences among the five groups, the number of response times at or below the over-all median of 7.6 sec. was obtained for each *S*. Using this measure, a test of the null hypothesis by simple analysis of variance resulted in an F less than one.

Although no significant differences were found, the curves suggest that some factor may have been operating in the 8-sec. group to retard learning in the runway. Because a guillotine door was present in both the Mowrer-Miller box and the runway, it is possible that generalization of a "waiting" or "resting" response occurred to this stimulus in the runway. It should be noted, however, that there is little difference in means at the end of acquisition. Using for each *S* the median response time of the last five trials, a test of the null hypothesis again yielded an F less than one.

Extinction.—To test extinction differences, the number of response times at or below the over-all median of

TABLE 1

MEANS AND SD'S OF EXTINCTION MEASURES IN FRUSTRATION-DRIVE TEST

Group	$\text{Log}_{10}(1+\text{WT})$		Trials to Extinction	
	Mean	*SD*	Mean	*SD*
0-sec.	2.19	1.01	22.16	18.51
2-sec.	2.78	.55	33.22	19.72
4-sec.	2.58	.92	35.95	19.49
8-sec.	2.14	1.11	17.47	15.08

73.3 sec. was obtained for each *S*. Extinction-score means are shown in Fig. 3, and the means and *SD*'s are presented in Table 2. A test of the linearity-of-regression hypothesis applied to the four transfer groups is significant beyond the .05 level ($F = 4.37$, $df = 1$ and 57). The control group does not differ significantly from the 0-sec. group ($t < 1$).

Because entering or not entering the goal box may be just as critical as the speed of the response, it was decided during the course of analysis that a second measure should be applied. This measure was the number of trials required to reach a criterion of two successive failures to enter the goal box. Trials-to-criterion means and *SD*'s are shown in Table 2. The test of the linearity-of-regression hypothesis applied to the four transfer groups is significant beyond the .025 level ($F = 6.04$, $df = 1$ and 57). As in the previous case, the control group did not differ significantly from the 0-sec. group ($t < 1$).

Discussion

Frustration-drive test.—The results of this experiment offer evidence for the occurrence of an increase in response vigor and resistance to extinction with the addition of blocking in avoidance extinction. Both of these findings are consistent with response expectations following from a frustration-drive interpretation.

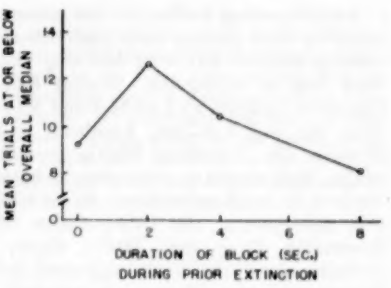

Fig. 3. Extinction in runway, measured by number of response times at or below the overall median, as a function of duration of blocking time in avoidance extinction.

As shown in Fig. 1 and Table 1, a reversal in both response vigor and resistance to extinction occurred when the duration of blocking exceeded a critical point. This reversal was predicted on the basis of the assumed dependence of frustration drive on initial level of drive intensity (**2**). Because secondary drive, fear, can be reduced by presenting the CS without shock, it was expected that the 8-sec. period of response blocking in the presence of the CS would result in a rapid diminution of initial drive. Thus less frustration drive was expected to develop in the 8-sec. group. A recent experiment by Page and Hall (**6**) reported similar results for prolonged blocking. Duration of blocking time is obviously a critical factor and needs to be examined under a variety of conditions utilizing both approach and avoidance responses.

When considering the shapes of the curves obtained for the various extinction measures, it should be remembered that the average latency (1.19 sec.) at the end of acquisition in avoidance learning is less than the minimum blocking time (2 sec.) used in this study. The overlap of .81 sec., in this case, may be called effective blocking time. It is possible, of course, that frustration drive could occur even if there is no period of overlap.

TABLE 2

Means and *SD*'s of Extinction Measures in Frustration–Stimulus Test

Group	Response Times at or below Median		Trials to Extinction	
	Mean	*SD*	Mean	*SD*
Control	9.27	4.56	11.07	5.95
0-sec.	9.27	4.32	9.53	5.33
2-sec.	12.73	5.44	15.00	6.50
4-sec.	10.20	6.36	13.13	6.86
8-sec.	8.20	6.22	9.40	6.95

An interesting finding in the present study is that groups with high wheel-turning activity following blocking also took long to extinguish. Similar findings were reported by Lambert and Solomon (4) and Lambert, Lambert, and Watson (3). Invoking Hull's concept of I_R, high response vigor should have resulted in rapid extinction. In view of this discrepancy, the incorporation of frustration drive into Hull's theory, perhaps along the lines suggested by Brown and Farber (2), appears indicated.

Frustration-stimulus test.—Evidence was obtained to support the hypothesis that response tendencies associated with the frustration state in one situation can transfer to other quite different frustrating situations. As predicted, slower extinction in the runway occurred for groups in which the avoidance response was blocked for short durations of time; the expected reversal of this trend for long durations of blocking time also occurred.

The prediction of response transfer was based on the hypothesis that a unique stimulus component, S_F, accompanies the occurrence of frustration drive. The basic similarity between the extinction and response-vigor curves in frustration-drive test and the extinction curves in frustration-stimulus test gives further support to the present theory. The transfer results also provide support for the assumed commonality in this study between blocking an avoidance response and not rewarding an approach response. Further exploration and definition of the class of stimuli to which these belong is required research in this area.

The question of precisely what responses or response tendencies were transferred is a difficult one. One possibility is that an increase in activity became associated with S_F in avoidance extinction, and that this response pattern was elicited during extinction in the runway. Whether or not the increased activity was goal directed requires further experimentation. Use of a choice situation instead of a straight alley in an experiment similar to the present one would provide an appropriate test.

Although less parsimonious, another possibility is that when S did not find food in the goal box, fractional antedating wheel-turning responses (r_G's) occurred. Referring to the Hullian model of serial response chaining, s's (in this case, S_F's) lead to fractional antedating goal reactions with their concomitant s_G's, which serve as secondary reinforcers. In this experiment, then, the accompanying s_G's would have reinforced the responses that led to the goal box, resulting in slower extinction for those Ss that had learned the appropriate association in avoidance extinction. It would be interesting in relation to this consideration to design a similar experiment with a wheel available in the goal box during extinction.

An alternative explanation of the results not dependent on the concept of S_F is that differential amounts of fear generalized from the experience with shock in frustration-drive test. If this were so, however, a difference should also have occurred between the 0-sec. group and the control group. Another factor tending to exclude fear as the critical variable is the fact that only slight differences were found in acquisition. With fear operating, acquistion should have been affected perhaps even more than extinction.

Some comment should be made concerning the possible effects of common external stimuli in the two situations (e.g., handling by the experimenter, lighting conditions, etc.). Such stimuli, of course, may have enhanced the response transfer that was of interest in this study; and because of their associations with fear on Day 1, these stimuli also may have contributed to criterion failure in the runway. It should be noted, with respect to both of these possibilities, that these stimuli were the same for all groups. The loss of Ss, which was only slightly differential among groups, should not affect the interpretation of the results. The alternative design of using a longer period of time between test days in order to reduce the loss of Ss was rejected because it was

felt that such spacing might attenuate the mediating effect of S_F.

Considering the results of this study, it may be argued that the relationships found in frustration-drive test can be satisfactorily interpreted solely in terms of S_F. That is, if S had developed learned reactions to frustrating events before serving in the experiment, it is possible that the response blocking on Day 1 merely reinstated or strengthened these reactions. Although this interpretation is possible, the different antecedent conditions defining frustration drive and frustration-drive stimulus in this experiment prescribe the use of two distinct concepts. It remains for further experimenting and theorizing to determine the explanatory power of the frustration-drive stimulus.

Summary

Two experimental situations were used to test some implications of the frustration drive and frustration-drive stimulus hypotheses.

1. The blocking of an avoidance response during extinction in a modified Mowrer-Miller box was used to determine the effect of the duration of response blocking on response vigor and resistance to extinction. The procedure during acquisition was the same for all groups; response-blocking times during extinction were 0, 2, 4, and 8 sec.

As predicted, a significant curvilinear relationship was obtained for both measures: An increase in response vigor and resistance to extinction occurred for short durations of blocking time, and a reversal in this trend occurred for prolonged blocking. The results were interpreted as a function of effective drive due to the interaction of frustration drive with secondary drive, fear; the reversal in the curve was attributed to a diminution in the intensity of the secondary drive.

2. On the day following avoidance training, the extinction of a running response in a straight alley was used to test for the transfer of response tendencies learned as a function of response blocking in avoidance extinction. Five groups were used: These included the four groups used on the previous day plus an added control group that had not received avoidance training. The procedure during acquisition and extinction was the same for all groups.

As predicted, a significant curvilinear relationship was obtained: An increase in resistance to extinction occurred for *S*s blocked for short durations of time in avoidance extinction; and a reversal in the curve commensurate with the relationships found in avoidance extinction was obtained. No significant difference occurred between the 0-sec. group and the control group. The results were attributed to the transfer of response tendencies that had become associated with frustration-drive stimulus.

References

1. Amsel, A., & Roussel, J. Motivational properties of frustration: I. Effect on a running response of the addition of frustration to the motivational complex. *J. exp. Psychol.*, 1952, **43**, 363–368.
2. Brown, J. S., & Farber, I. E. Emotions conceptualized as intervening variables—with suggestions toward a theory of frustration. *Psychol. Bull.*, 1951, **48**, 465–495.
3. Lambert, W. W., Lambert, E. C., & Watson, P. D. Acquisition and extinction of an instrumental response sequence in the token-reward situation. *J. exp. Psychol.*, 1953, **45**, 321–326.
4. Lambert, W. W., & Solomon, R. L. Extinction of a running response as a function of distance of block point from the goal. *J. comp. physiol. Psychol.*, 1952, **45**, 269–279.
5. Mowrer, O. H., & Miller, N.E. A multipurpose learning-demonstration apparatus. *J. exp. Psychol.*, 1942, **31**, 163–170.
6. Page, H. A., & Hall, J. F. Experimental extinction as a function of the prevention of a response. *J. comp. physiol. Psychol.*, 1953, **46**, 33–34.
7. Rohrer, J. H. A motivational state resulting from non-reward. *J. comp. physiol. Psychol.*, 1949, **42**, 476–485.
8. Sheffield, V. F. Extinction as a function of partial reinforcement and distribution of practice. *J. exp. Psychol.*, 1949, **39**, 511–526.
9. Snedecor, G. W. *Statistical methods.* Ames, Iowa: Iowa State Coll. Press, 1946.
10. Stanley, W. C. Extinction as a function of the spacing of extinction trials. *J. exp. Psychol.*, 1952, **43**, 249–260.
11. Thompson, W. R. On a criterion for the rejection of observations and the distribution of the ratio of deviation to sample standard deviation. *Ann. math. statist.*, 1935, **6**, 214–219.

(Received May 14, 1956)

Journal of Experimental Psychology
Vol. 54, No. 2, 1957

SOME FACTORS AFFECTING PERFORMANCE ON A WORD-FORMATION PROBLEM[1]

WILLIAM F. BATTIG[2]

Stanford University

Several recent reviewers (**6**, **10**, **14**) have concluded that progress in research on complex types of human behavior (i.e., thinking, reasoning, problem solving) has been discouragingly slow, despite a steady increase in quantity. It has been suggested (**10**, **13**) that this lack of progress can be at least partially attributed to the failure to develop standardized experimental tasks for such research, comparable to such laboratory standbys in more highly developed areas of psychology as the nonsense syllable, Skinner box, etc. Ray (**10**) lists 29 tasks used in problem-solving research. However, few of these have been used in more than one or two experiments, and none has been subjected to the kind of extensive analysis necessary even to specify the characteristics of the task itself.

Ray states that useful tasks for problem-solving research should fulfill the criteria of (*a*) a longer and more sensitive scoring continuum than the common "pass-fail" variety, and (*b*) providing a maximum amount of knowledge of what *S* is doing. Suggested ways of meeting the second criterion include limiting the number of possible hypotheses, allowing only one possible solution, and maximizing the amount of overt behavior involved. To these criteria might be added (*c*) sufficient interest so that *S*s will be easily motivated to work at the problems, and (*d*) some relationship to problems confronted in everyday life.

The general category of verbal problems appears to fulfill most of these criteria. The great popularity of crossword puzzles, Scrabble, and the like attest to the practical interest in such tasks. In addition, the number of elements involved is finite and identifiable, and there has already been considerable progress in the quantification of the characteristics of verbal material (**7**). Despite these advantages, such problems have seen little service in the psychological laboratory. Anagrams have been used by some investigators, but this task gives little evidence about the psychological processes involved, and is relatively insensitive as a measuring instrument, as it can be scored only in terms of pass-fail or time to completion.

A type of verbal problem which seems to satisfy most of the proposed criteria is one which has been termed *word formation*, originally suggested by the formerly popular television program of the American Broadcasting Co., *Down You Go*. This task involves the identification of common phrases, knowing only the number of letters in each word, by guessing letters of the alphabet. Each time a letter is guessed, *S* is told the positions the letter holds, if the letter occurs in the phrase. The objective is to identify the phrase by naming as few letters as possible.

The present paper reports the results of an exploratory study of this

[1] This research was done while the author was a post-doctoral fellow of the National Science Foundation. Appreciation is due to Dr. E. R. Hilgard, fellowship adviser, for his guidance in the conduct of this research, and for a critical reading of the manuscript.

[2] Now at the University of Virginia.

TABLE 1

WORDS USED

Frequency of Usage	Length				
	4	5	6	7	8
1	sure(9)*	fruit(11)	basket(14)	certain(9)	neighbor(16)
4	pang(17)	torch(13)	combat(13)	scholar(15)	creation(9)
7	clod(13)	opium(12)	visage(17)	honesty(16)	dismount(11)
10	dual(11)	resin(8)	malign(15)	inflate(15)	demolish(15)

* *Expected scores* based on letter preferences in guessing.

task at the level of words instead of phrases, with the primary purpose of investigating the effects of variation in the length (number of letters) and frequency of usage of words. Probably more significant was the emergence of several variables associated with individual differences in performance on the task.

METHOD

Forty introductory psychology students at Stanford University served as *Ss* in this experiment. A set of 20 words was selected from Thorndike's list (12) so that all combinations of five lengths (4, 5, 6, 7, and 8 letters) and four frequencies of usage (from the first, fourth, seventh, and tenth thousand words ranked in order of frequency of occurrence in English text) were represented once. Selection of the word within each category was random. The same 20 words were used for all *Ss*, and are given in Table 1. Twenty different sequences of presentation were provided by adopting a 20 × 20 latin-square design, so that each word appeared in each position of the sequence an equal number of times. The same square was replicated twice, therefore two *Ss* underwent each of the word sequences.

At the outset of the experiment, all *Ss* were read the following instructions:

"In this experiment we will play a kind of word game, in which your task will be to find a particular word by naming letters of the alphabet until you have named all the letters in the word. The object is to find out what the word is by naming as few letters as possible. The fewer letters you require, the better your score will be. Each time you name a letter, I'll tell you if it is in the word, and if so, what position it holds, that is, if it is the first letter, second letter, and so on. You are to continue naming letters until you have named all the letters in the word, and your score will be the number of letters required to do this. I am also going to time you, but I am mainly interested in number of letters rather than time.

"To begin with, obviously you won't be able to do much more than guess, although you know that some letters of the alphabet are used more often than others, and you will be more likely to find a correct letter by naming the more common ones. However, as you get more and more letters in the word, you should begin to get an idea of what the word is, and you can do a lot better than just guessing by making use of this information. For example, if you know that *und-* are the first three letters of a five-letter word, you may recognize that the word could be *under*, and can complete it by naming *e* and *r*. This is the object of the game, to learn to make use of this information so that you can get the word by naming as few letters as possible.

"The words which we will use have all been selected according to certain rules. First, no word contains any one letter more than once, so that words like queen or toot are eliminated. There will be no proper names, and all the words are taken from the 10,000 most common words in the English language, so they should all be reasonably familiar to you. Do you have any questions?"

All *Ss* were provided with worksheets consisting of spaces to write down the correct letters as they were named, along wth 20 lists of letters in alphabetical order, makinig it possible to cross out or check letters already guessed.

Before each word, *S* was told the number of letters it contained. Each time a letter was guessed, *E* responded "No" if the letter was not in the word, but gave the position (first letter, second letter, etc.) if the letter was in the word. This was continued until *S* had named all the letters in the word. The *E* recorded the order in which letters were guessed and the time re-

quired to complete each word. No knowledge of results was given until the end of the experiment.

Results

Two direct measures of performance were obtained, a trials score (number of letters guessed for each word), and the time required to solve each word. In addition, a recognition score was obtained, consisting of the number of correct letters guessed consecutively without error up to, but not including, the final correct letter which was named.

Word differences.—The means of the three measures for each word are presented in Table 2. Sizeable differences in performance between words are apparent, much of which cannot be accounted for in terms of the manipulated variables of word length and frequency of usage. The effect of length appears to be curvilinear, except for the recognition measure. For the other measures, performance first declines with increased length, but subsequently improves. As frequency of usage decreases, performance becomes poorer, especially with the time score.

All measures were subjected to identical analyses of variance for the latin-square design (**3**, **4**) combined with Grant's extension (**5**) of Alexander's trend test (**1**). These are summarized in Table 3. The results for the three measures are similar, and will be discussed together.

TABLE 2

Mean Performance Measures on Each Word

Frequency	Length					Total
	4	5	6	7	8	
			Trials			
1	12.83	13.80	14.60	11.50	14.15	13.38
4	16.48	15.25	13.60	16.00	12.20	14.71
7	12.83	14.43	17.18	15.85	13.75	14.81
10	13.50	13.13	16.25	14.38	14.03	14.26
Total	13.91	14.15	15.41	14.43	13.53	
			Time (in Seconds)			
1	56.25	68.65	73.55	53.40	62.78	62.93
4	85.25	75.88	73.18	84.88	66.33	77.10
7	70.33	77.18	92.35	80.75	69.43	78.01
10	64.20	68.90	88.33	96.43	77.25	79.02
Total	69.01	72.65	81.85	78.86	68.94	
			Recognition			
1	0.60	1.08	1.43	1.90	2.58	1.52
4	0.23	0.95	0.98	1.55	2.90	1.32
7	0.90	1.23	1.00	1.90	2.10	1.43
10	0.70	0.83	1.38	1.03	2.08	1.20
Total	0.61	1.02	1.19	1.59	2.41	

TABLE 3

SUMMARY OF ANALYSES OF VARIANCE

Source	Error Term	*df*	Trials		Time		Recognition	
			MS	*F*	*MS*	*F*	*MS*	*F*
1. Between *Ss*	9	39	78.05	4.11**	13826	9.44**	3.09	2.23**
2. Sequences	3	19	79.89	1.05	14547	1.11	2.09	0.52
3. *Ss* within Seq.	—	20	76.30	—	13142	—	4.04	—
4. Words	9	19	90.26	4.75**	5196	3.55**	18.92	13.69**
5. Length (L)	9	4	80.25	4.22**	5490	3.75**	74.98	18.56**
a. Linear	9a	1	3.52	0.24	593	0.43	280.56	126.38**
b. Quadratic	9b	1	233.36	9.71**	17663	13.49**	12.30	7.94**
6. Frequency (F)	9	3	85.00	4.47**	11549	7.89**	3.69	0.91
a. Linear	9c	1	75.08	4.69*	24197	17.52**	3.53	2.87
7. L × F	9	12	94.92	5.00**	3511	2.40**	4.04	2.92**
8. Order	9	19	24.16	1.27	2690	1.84*	0.85	0.62
a. Linear	9	1	38.88	2.05	10852	7.41**	1.87	1.38
9. Residual error	—	722	19.00	—	1464	—	1.38	—
a. Between ind. linear (length)		39	14.24	—	1393	—	2.22	—
b. Between ind. quadratics (length)		39	24.03	—	1309	—	1.55	—
c. Between ind. linear (frequency)		39	16.02	—	1381	—	1.23	—

* Significant at .05 level.
** Significant at .01 level.

The replicated latin-square design provides two error estimates, one reflecting the confounding of interactions inherent in the particular square which was used, the other a residual variability within the cells of the square. These did not differ significantly for any of the three measures, and have been pooled to provide a single more reliable error estimate (**4**).

Differences between words were significant for all measures. The design permits the isolation of the effects of word length, frequency of usage, and the interaction between them. The effects of these variables are complex and not consistent for the various measures of performance.

The length term is significant in all cases when tested against the residual error term. Application of orthogonal polynomials (**5**) reveals a significant linear component of the curves relating word length to performance only for the recognition score. Only the quadratic component is significant for the other measures, due to the decline in performance as length increases from 4 to 6 letters, followed by an improvement for 7 and 8 letters. The significant linear component for the recognition score can be attributed to an artifact, since the maximum possible recognition score increases linearly with increased length. This suggests that a certain proportion of the letters must be known for a word to be recognized. The significant quadratic component of the recognition curve reflects the disproportionate increase in this score as length increases, particularly the marked increase from 7 to 8 letters.

Frequency of word usage is significant, except for the recognition score. The linear component of frequency is significant for both the trials and time measures, indicating that performance declines as frequency of usage decreases.

The sizeable and significant inter-

action for all measures prevents any generalization about the effects of word length and frequency beyond the specific words used in the present study. This also suggests that other variables may account for word differences, particularly with the trials score where the interaction mean square exceeds that for either length or frequency.

The Between-*Ss* term is significant for all measures, showing reliable individual differences in performance. In no case was the Sequences term significant when tested against the appropriate *Ss* within Sequences error term. This means that over-all performance was not affected by the particular sequence in which the words were given.

The Order term reflects any change in performance with practice, and its linear component provides a test for consistent improvement or decrement with practice. These are significant only for the time score. It appears that the rate of naming letters increases with practice, but there is no change in the number of letters required to solve the word.

Examination of the data indicated that word differences might be partially accounted for in terms of the letters they contain. This task places a premium on guessing, so that a word containing letters likely to be guessed early should be solved in fewer trials than one which does not. Prior data on the order in which letters are guessed was lacking, so this variable was not controlled in selecting the words. However, a tabulation of the number of times each letter was guessed first, second, etc. over all words for each *S* provided an estimate of this effect. A *letter preference score* was obtained by assigning each first guess a score of 25, each second guess 24, etc. so that if all 26 letters were guessed before the word was solved, the 26th letter had a score of zero. Any letter which was not guessed on a particular word was also given a score of zero. This score gives an index of the relative order in which letters are guessed, so that the higher the score, the more likely the letter was to be guessed early. The letters were ranked in terms of this measure from 1 to 26, and the expected number of trials for each word was calculated under the assumption that letters would be guessed in this particular order. Since the results on the recognition measure indicated that *S* was usually able to recognize the word when all but one letter were known, this expected score did not include the least preferred letter in the word. The procedure may be clarified by an example, using the word "sure." The letter *e* ranked second in preference, *s* was sixth, *r* was eighth, and *u* was tenth. The expected score is nine (the number of guesses required to get all but one letter, plus one for the last letter). These expected scores are given in parentheses after each word in Table 1.

Despite the crudeness of this measure, and its failure to take into account the effect of length on the recognition score, a correlation of .821 was obtained between the actual scores for the 20 words, and the calculated expected scores. This indicates that at least two-thirds of the variance between words on the trials measure can be attributed to differences in the letters they contain.

Individual differences.—An extensive analysis was undertaken to attempt to identify factors associated with individual differences in performance. Comparison of the upper and lower quartiles on the trials measure revealed significant relationships to several factors. The trials score was selected for this differentiation because *Ss* were instructed to work for this score. The upper quartile consisted of 10 *Ss* with mean trial scores of 13 or below, while the lower quartile included 11 *Ss* with mean scores of 15 or above.

The order in which letters were guessed appeared to be a likely source of performance differences. Analysis of variance of the letter preference scores gave a significant interaction between letters and quartiles ($F = 7.10$, 25 and 475 *df*). Closer examination indicated that the upper quartile letter preference scores corresponded more closely to the actual frequencies with which letters are used in the English language, while the lower quartile tended to follow alphabetical order. The letters *r*, *s*, and *t* ranked higher for the upper quartile, while *b*, *c*, and *d* had higher scores for the lower quartile. The statistical significance of this effect was tested

with an adaptation of the Alexander trend test (1), by analyzing the trends when letters were ordered alphabetically, and again when ordered according to frequency of usage (9). The slopes of the curves for the upper and lower quartiles differed significantly according to both methods of ordering. The lower quartile slope was significantly steeper ($F = 21.80$, 1 and 19 *df*) for alphabetical order, but the upper quartile slope was steeper ($F = 14.80$, 1 and 19 *df*) when ordered according to frequency of usage. This analysis demonstrates that *S*s who perform well on this task tend to guess letters according to the frequency with which they occur in the English language, while those who perform poorly are more closely bound by alphabetical order.

Previous work on problem solving has made much of the fact that efficiency is interfered with by a lack of flexibility, or failure to vary the approach to the problem (6). This would be manifested in the present task by a tendency to guess letters in the same order from word to word. An index of flexibility was obtained from the letter preference scores by calculating the proportion of the variance of these scores which could be accounted for by differences between letters of the alphabet. This measure is similar to the eta correlation coefficient (8), and was calculated for each *S* in the upper and lower quartiles by determining the ratio of the sum of squares between letters to the total sum of squares. The magnitude of this value is directly related to the consistency of the order in which letters are guessed from word to word, reaching a maximum of one if the identical order is used for all words, and a minimum of zero if the order is random. This index did not differ significantly for the upper and lower quartiles ($t = .57$), but the direction of the difference is consistent with previous work, as the upper quartile index is below the lower quartile.

The nature of the word-formation task is such that the most efficient approach would be to guess the most common letters first. After some letters in the word are known, subsequent guesses should be based on the knowledge of these letters rather than solely on frequency of usage. This suggests that good performance should be associated with consistency in guessing the first few letters, followed by a shift to increased variability. Such a comparison requires a measure which differentiates variability from beginning to end of the sequence of guesses. Letter preferences have been shown to differ for the upper and lower quartiles, eliminating analysis in terms of specific letters. Instead, a comparison of the means and *SD*'s of the upper and lower quartiles was made for the most preferred letter, second most preferred letter, etc., regardless of what the letters were. Increased variability is indicated by either lower means, higher *SD*'s, or both. Such an analysis is limited because not all letters were guessed for each word. However, there were few cases where a word was solved in less than 10 trials, so the mean letter-preference scores and *SD*'s were calculated for the 10 most preferred letters for the upper and lower quartiles. These are presented in Fig. 1.

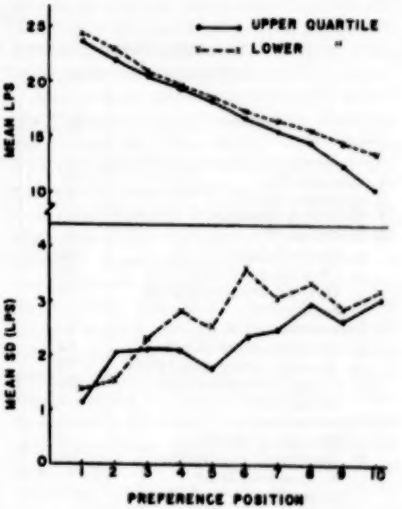

FIG. 1. Means and *SD*'s of letter-preference scores (LPS) for 10 most preferred letters for *S*s in upper and lower quartiles on trials score.

The summary of an extended trend analysis of these data is presented in Table 4. A significant difference between Quartiles is shown for the means measure, as is a Quartiles-by-Positions interaction. The linear component of the interaction is not significant, but the quadratic component is. This can be accounted for by the increasing separation of the curves over the last few positions. When only the last six points are considered, there is a significant difference between group slopes ($F = 8.18$).

For the *SD* measure, neither the difference between quartiles nor the interaction is significant, but the quadratic component of the interaction is significant at about the 2% level. Further analysis showed a significant quadratic component for the upper quartile curve, but not for the lower quartile. Figure 1 shows that this difference can be attributed to a smaller change in *SD* by the upper quartile over the first few letters, followed by a relatively rapid increase, as contrasted to the lower quartile.

TABLE 4

EXTENDED TREND ANALYSIS OF 10 MOST PREFERRED LETTERS

Source	Error Term	*df*	Means		*SD*'s	
			MS	*F*	*MS*	*F*
1. Between *Ss*	6	20	2765	4.05**	58034	7.53**
2. Quartiles (Q)	3	1	25671	16.47**	83558	1.47
3. *Ss* within Quartiles	—	19	1559	—	56691	—
4. Positions (P)	6	9	130966	191.75**	79655	10.34**
a. Over-all slope	6a	1	1173641	484.77**	559909	72.35**
b. Over-all quadratic	6b	1	432	0.41	74432	9.62**
c. Over-all deviations from estimation	6c	7	660	1.72	11794	1.54
5. Interaction (Q × P)	6	9	1948	2.85**	11323	1.47
a. Between-group slopes	6a	1	9929	4.10	4505	0.50
b. Between-group quadratics	6b	1	6698	6.38*	45526	6.69*
c. Group deviations from estimation	6c	7	129	0.33	7411	0.97
6. Residual error	—	171	683	—	7707	—
a. Between-*S* slopes	—	19	2421	—	8967	—
b. Between-*S* quadratics	—	19	1050	—	6806	—
c. *S* deviations from estimation	—	133	382	—	7663	—

* Significant at .05 level.
** Significant at .01 level.

Both measures show for the upper quartile little change in consistency of guessing the first few letters, followed by a sharp increase in variability. In contrast, the lower quartile choices decrease in consistency in approximately linear fashion throughout the curve.[3]

Comparison of the trials scores of the upper and lower quartiles on the 20 words revealed the presence of a significant Quartiles-by-Words interaction ($F = 2.50$, 19 and 361 *df*). The Quartiles-by-Length interaction was significant ($F = 3.94$, 4 and 361 *df*), but the Quartiles-by-Frequencies term was not ($F = .36$). The triple interaction (Quartiles by Length by Frequency) was also significant ($F = 2.55$, 12 and 361 *df*), indicating that part of the differential performance of the quartiles could not be accounted for in terms of length or frequency of usage. The extended trend test showed that the Quartiles-by-Length interaction was due to the presence of the quadratic relation between the trials score and Length only for the lower quartile. The between-group quadratics terms was the only significant interaction component ($F = 11.87$, 1 and 19 *df*), and further analysis revealed a significant quadratic component for the lower quartile, but none for the upper quartile.

A comparison of learning effects for the two quartiles showed slightly more improvement with practice for the lower quartile, but the difference between group slopes was nonsignificant ($F = 3.30$, 1 and 361 *df*).

A time-per-trial score, obtained by dividing the total time for each word by the number of trials, was found to be significantly greater for the upper than the lower quartile ($t = 2.45$, $P < .05$), indicating that the upper quartile took more time per guess than the lower quartile. This is in accord with a slight insignificant negative correlation obtained between the time and trial measures ($r = -.120$).

Verbal and mathematical aptitude scores on examinations of the Educational Testing Service were available for all *Ss*. The trials measure was significantly correlated with verbal aptitude scores ($r = .468$), but the correlation with mathematical aptitude scores was not significant ($r = .206$).

[3] Similar analyses of individual differences have since been performed on data from another experiment using this task. The results obtained in the present paper were generally confirmed, with the single exception of the *SD* measure. This was found to be significantly lower for the upper than for the lower quartile, and there was no evidence for a difference in the shape of the two *SD* curves such as that shown in Fig. 1. Extension to 15 preference positions still gave no indication that the curves were converging. These results require some qualification in conclusions about the *SD* measure, but do not basically alter the interpretation of the differences between the upper and lower quartiles.

Discussion

Conclusions about the effects of word length and frequency of usage are severely limited by the small number of words, along with evidence for the strong influence on performance of specific letters contained in the words. Nevertheless, the failure to find a monotonic increase in the trials score with increased length is surprising, as such a relationship should obtain on a strict probability basis. Under the unrealistic but simplifying assumptions of independence of successive guesses and equal probability of all letters being guessed, the trials score should increase from 21.6 for 4 letters to 24.0 for 8 letters. However, the probability of any particular letter being correct also increases linearly with word length. Previous studies (**2, 11**) have shown a rapid increase in the probability of getting the correct letter in English context as the number of known preceding letters increases. In the language of information theory, the redundancy increases and the transmitted information declines as more letters in the word are known. If this change is sufficiently rapid, the increased probability that initial guesses will be correct may lead to a lower trials score as words increase in length. The finding of Zipf (**15**) that the number of words of a given length declines with increased length, so that there are fewer words to choose from as length increases, provides another plausible explanation for this result.

The present data lack the precision necessary for a mathematical specification of these relationships, particularly since they differ for the upper and lower quartiles. Only a suggestion is justified that trials to solution increase with word length up to a certain point, and then level off or decrease. The exact nature of this effect is a topic for future research.

The finding that *S*s with low trials scores show greater correspondence of their letter preference scores to the frequencies with which letters are used in English is of considerable interest. This could be due to more knowledge about which letters are most common, or the result of a greater tendency to guess letters on the basis of actual words or common letter combinations.[4]

No learning effect was found for either the trials or recognition scores. Practice conditions in this study were highly massed, which may have obscured the presence of any learning.

Several lines of evidence, particularly the data of Fig. 1, suggest that the most efficient approach to this problem is somewhat as follows. One should begin by naming the most commonly used letters until a certain number of correct letters are known. Thereafter, further guesses should be based on known words which would fit, or common letter combinations. The point of this transition will vary from word to word, and also may differ for individuals. The most common approach was to guess the vowels first, usually in alphabetical order. This procedure is relatively effective, but the guessing of *u* is efficient only if the number and distribution of other vowels suggest its presence. The difference in preference scores for *u* between the upper (19.6) and lower (22.2) quartiles indicate that this factor may have contributed to individual differences.

A brief comment is in order about the generalization that problem-solving efficiency is enhanced by flexibility of approach. When the presence of several possible responses and the unlikelihood that the first response will be correct are made defining properties of problem solving (**6, 10**), it would be surprising if this were not the case. The present study is concerned with a problem where it is not, as efficiency was improved by a high degree of consistency in the initial stages

[4] Subjects were asked to rank the letters of the alphabet in terms of frequency of usage at the end of a later experiment with this task. A measure of the rank correlation between this ordering and the actual ordering in terms of frequency of usage was significantly correlated with the trials measure ($r = .308$), suggesting that knowledge of the frequency with which letters are used is related to performance on this task.

of the various problems. Other similar examples could be cited, and only a very arbitrary distinction could exclude these from the area of problem solving. The ability to vary one's approach according to the requirements of the problem may be associated with problem-solving efficiency, but this restatement deprives the generalization of much of its psychological significance, as it emphasizes the dependency on the nature of the problem. This points up the dangers involved in overgeneralization, and speaks for the need for more of the extensive analyses of the factors involved in specific problems which the present approach exemplifies.

In conclusion, it can be said that the word-formation problem at least shows promise of becoming a useful tool for problem-solving research. Its major shortcoming is clearly the dependence of performance measures on the particular words which are used. Although this is a variable of considerable interest, failure to take it into account and to control it severely limits the use of the method for the investigation of other variables. An extensive calibration, based on the empirical determination of word difficulties, appears necessary to provide materials sufficiently standardized to give the technique general applicability.

Summary

A word-formation problem was investigated, in which *S*'s task is to identify words by naming letters of the alphabet until he gets all of the letters in the word. Forty *S*s solved 20 words representing all combinations of five lengths (four through eight letters) and four frequencies of usage, presented in counterbalanced orders through a latin-square design. Results showed significant variability in performance between words, much of which could be attributed to the specific letters contained in the words. Limited evidence for the effects of length and frequency of usage was found, and there was little or no evidence for learning on the task. Comparison of *S*s in the upper and lower quartiles in performance revealed: (*a*) the upper quartile tended to guess letters in the order of their frequency of usage, while the lower quartile was more bound by alphabetical order; (*b*) no over-all difference in variability of approach was found, but the upper quartile was relatively more consistent in its first few guesses, and became increasingly less consistent on subsequent guesses, as compared with the lower quartile; and (*c*) the upper quartile took significantly more time per guess than the lower quartile.

References

1. Alexander, H. W. A general test for trend. *Psychol. Bull.*, 1946, **43**, 533–557.
2. Burton, N. G., & Licklider, J. C. R. Long-range constraints in the statistical structure of printed English. *Amer. J. Psychol.*, 1955, **68**, 650–653.
3. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
4. Grant, D. A. The latin square principle in the design and analysis of psychological experiments. *Psychol. Bull.*, 1948, **45**, 427–442.
5. Grant, D. A. Analysis of variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **53**, 141–154.
6. Johnson, D. M. *The psychology of thought and judgment.* New York: Harper, 1955.
7. Miller, G. A. *Language and communication.* New York: McGraw-Hill, 1951.
8. McNemar, Q. *Psychological statistics.* New York: Wiley, 1949.
9. Pratt, F. *Secret and urgent.* Garden City, N. Y.: Blue Ribbon Books, 1942.
10. Ray, W. S. Complex tasks for use in human problem-solving research. *Psychol. Bull.*, 1955, **52**, 134–149.
11. Shannon, C. E. Prediction and entropy in printed English. *Bell Syst. Tech. J.*, 1951, **30**, 50–64.
12. Thorndike, E. L. *Thorndike-Century senior dictionary.* New York: Scott, Foresman, 1941.
13. Underwood, B. J., & Richardson, J. Some verbal materials for the study of concept formation. *Psychol. Bull.*, 1956, **53**, 84–95.
14. Woodworth, R. S., & Schlosberg, H. *Experimental psychology.* (Rev. ed.). New York: Holt, 1954.
15. Zipf, G. K. *Human behavior and the principle of least effort.* Cambridge, Mass.: Addison-Wesley, 1949.

(Received June 5, 1956)

Journal of Experimental Psychology
Vol. 54, No. 2, 1957

EFFECT OF ADDITION OF IRRELEVANT VERBAL CUES ON PERCEPTUAL-MOTOR LEARNING

HARRY W. BRAUN AND A. W. BENDIG[1]

University of Pittsburgh

The investigation of the role of verbal cues in motor learning has been characterized in the main by two methods. In one case, *Ss* are given verbal pretaining on stimulus and response words which are relevant or irrelevant to the stimuli and responses of the motor task, and the effect of such pretraining on subsequent performance in the motor task is observed (**2, 7**). The results of studies using this method have not been consistent. In the second method, no pretraining is given, but verbal cues are added to the motor task during acquisition of the skill in question and the effect of this addition on performance is measured (**1**). In the few studies which employed this latter method, the verbal cues are relevant in that they are "verbal descriptions of the stimulus, the response, or both" (**1**, p. 371). Inconsistent results have also been obtained with this method. While maze learning is facilitated by relevant verbalization (**9**), the acquisition of a lever-positioning skill is not affected by the addition of relevant verbal cues (**1**,

The primary purpose of the present investigation was to determine the effect of the addition of irrelevant verbal cues on the acquisition of a perceptual-motor skill. The cues were high- and low-meaningful terms drawn from Noble's *m*-scale (**8**). A secondary aim of the study was to ascertain the relationship of several psychometric measures to performance on the task.

[1] The authors acknowledge the assistance of Mrs. Theresa Miller and Mrs. LaVene Taber in running the experimental groups.

METHOD

Apparatus.—The apparatus was designed and constructed by J. F. Pierce and may be broadly classified as an electromaze or as a multiple-choice apparatus. Mechanically, it is very flexible in the sense that a multiplicity of problems may be set up, ranging from a simple 2-choice situation to a complex 11-choice problem. On *E*'s side of the instrument are 11 lights with red shields below which are 11 double-pole double-throw switches. These switches permit *E* to render inoperable any light or any combination of lights. Below the switches are 11 rotary multiple-contact switches with 11 contacts each. These switches permit *E* to set up any pattern of lights with any number (up to 11) of lights in the pattern. On *S*'s side of the apparatus are 11 lights, also in a horizontal row, which correspond to the lights of *E*. Below these lights are 11 momentary contact switches. The choices of *S* must be made in a given pattern or sequence for the solution of the problem that is posed. The *S* is required to identify the pattern and to press the momentary contact switch for each step in the sequence. A correct choice results in the closure of an electrical current through the light for that switch. In this manner *S* is informed of the success or failure of his choice. Furthermore, an incorrect choice results in the opening of all previously closed light circuits and a buzz from the apparatus. Not only is *S* informed of his failure but he must begin a new trial after each incorrect choice, continuing in this manner until some criterion of performance is attained.

Subjects.—Fifty-five male and female undergraduate students enrolled in psychology classes at the University of Pittsburgh.

Design.—The *Ss* were divided into five groups of 11 each. Group 1 (Control) learned a given 11-button sequence by any means at their disposal. Group 2 learned the sequence and vocalized a high-meaningful word which was placed above each switch when each switch was pressed. Group 3 vocalized a low-meaningful word above each button in learning the sequence. In Group 4 the same high-meaningful words

were present above each switch but *Ss* were not required to vocalize them. A similar condition prevailed for Group 5 except that low-meaningful words were posted.

Verbal cues.—The 11 high-meaningful words were: captain, zebra, garment, dinner, village, money, insect, jewel, army, heaven, office. They were drawn from Noble's list (**8**) and their mean *m* value was 7.82. The 11 low-meaningful cues were: volvap, neglan, latuk, tarop, quipson, gokem, kupod, polef, zumap, nostaw, and sagrok. They also were taken from Noble and had a mean *m* value of 1.28.

Problem sequences.—Eleven random sequences of 11 button (switch) pushes were used, a different sequence being required of each *S* of any of the five groups, with each sequence representing one row of an 11 × 11 latin-square design. For each group the rows constituted the problem sequences (and individual *Ss*). The columns represented the ordinal position of the sequential responses, while the cell entries were the individual button switches. The *S*'s task was to learn the sequence of 11 button pushes which would light all the lamps to a criterion of two errorless trials. Each *S* used his own pacing rate in pressing the button switches.

Results

The variances of the raw trials-to-criterion scores within the five groups of *Ss* were tested for homogeneity by Cochran's method (**4**, p. 147) and the resulting value of .426 was significant at the .05 level. However, the same test applied to the logarithms of the trial scores yielded a value (.255) that did not approach significance at the .05 level. While the raw trials scores were positively skewed, the distribution of the log trials scores was approximately normal.

TABLE 1

Analysis of Variance of Transformed (log) Trials-to-Criterion Scores

Source of Variation	*df*	*MS*	*F*
Groups (G)	4	.0818	3.60*
Control vs. Verbal Cues (C)	1	.2400	10.57**
Vocalization (V)	1	.0743	3.27
Meaning (M)	1	.0014	.06
V × M	1	.0117	.52
Sequences (S)	10	.0321	1.41
G × S	40	.0227	
S × C	10	.0282	.64
S × V	10	.0064	.15
S × M	10	.0121	.27
S × V × M	10	.0441	
Total	54		

* $P = .05$.
** $P = .01$.

The log trials scores were subjected to a two-criterion analysis of variance and the results are shown in Table 1. The Groups × Sequences error mean square was divided into its four components and these were found to be homogeneous. Consequently, these pooled components, or the Groups × Sequences mean square, were used as the error term in testing the significance of the main effects. The differences among the groups were significant at the .05 level. The significance of the group differences was also tested by the conventional median test and the rank test of Kruskal and Wallis (**6**). The median test yielded a chi square of 10.62 which, with 4 *df*, is significant at the .05 level, while the rank test (corrected for ties) gave a chi square of 11.00 which is also significant at the .05 level. A more refined analysis of the intergroup differences indicated that the significance was attributable to the difference between the control group (Group 1) and the combined verbal cues groups (Groups 2, 3, 4, and 5), which was significant at the .01 level, and, possibly, to the difference between the aloud (Groups 2 and 3) and silent (Groups 4 and 5) vocalization groups (.10 level). The difference between the high- and low-meaningfulness groups and the interaction of vocalization and meaningfulness were not significant. Table 2 gives the means and *SD*'s of both raw trials and log trials scores for each of the five groups. It can be seen that the control group was clearly inferior

in performance compared to the verbal cues groups, while the aloud vocalization groups were slightly superior to the silent vocalization groups.

Scores on three psychometric measures were available for 51 of the *Ss*: a 20-item vocabulary scale drawn from the Cooperative Vocabulary Test, a 20-item version of Taylor's Manifest Anxiety Scale (**3**), and Edwards' 28-item Need Achievement Scale which is included in the *Personal Preference Schedule* (**5**). Analyses of variance of stanine scores from these tests showed the aloud-silent vocalization groups to be significantly different (.05 level) on the vocabulary scale and the high-low meaningfulness groups to be significantly different (.05 level) on the MAS. Analyses of covariance were computed, adjusting the log trials scores for each single test and each combination of tests, but none of the resulting analyses of variance of the adjusted log trials scores changed the conclusions drawn from the analysis reported in Table 1. The average within-group correlations between the three tests and log trials scores were: Vocabulary, $r = -.21$; MAS, $r = .00$; and Need Achievement, $r = -.38$. The correlation between Need Achievement and the learning measure (log trials) was significant at the .01 level, but neither of the other two coefficients was significant at the .05 level.

TABLE 2

NUMBER OF TRIALS TO CRITERION

($N = 11$ each group)

Group	Raw Scores		Log Scores	
	Mean	*SD*	*Mean*	*SD*
1. Control	68.2	27.5	1.80	.17
2. High-Aloud	39.8	11.6	1.58	.12
3. Low-Aloud	41.8	12.8	1.60	.12
4. High-Silent	53.6	19.3	1.70	.17
5. Low-Silent	48.3	18.6	1.65	.16

DISCUSSION

The results of this study show that the development of a specific perceptual-motor skill can be facilitated by the addition of irrelevant verbal cues. As such, they suggest an extension of the hypothesis attributed to Brogden "That the acquisition of a motor skill will be facilitated to the extent to which verbal descriptions of the stimulus, the response, or both, can be utilized" (**1**, p. 371) to include a similar facilitating effect for irrelevant verbal cues.

As transfer from verbal pretraining to subsequent motor performance has been shown to be a function of the complexity of the task (**2**), so the facilitation of perceptual-motor skill development by verbal cues may be related to the nature or complexity of the task. The perceptual-motor skill which was developed by *Ss* in this study may be regarded as an instance of a relatively homogeneous chain because of the high degree of similarity between the discriminative stimuli as well as between the responses in the chain. The addition of verbal cues in the form of printed irrelevant words placed above the response buttons is interpreted as facilitating the acquisition of the skill in question by increasing both stimulus and response discriminability. This interpretation is supported by the tendency for the groups which vocalized the verbal cues aloud to be superior to the groups which vocalized the cues silently. It must be admitted, however, that the failure to obtain differential effects for high- and low-meaningful verbal cues does not seem to support this interpretation. On the basis of the interpretation of the results which has been advanced, the prediction would follow that the facilitating effect of irrelevant verbal cues would not be as great in the case of a heterogeneous perceptual-motor chain where less similarity exists between both stimuli and responses.

The results using the three psychometric measures suggest that individual performance on our complex learning task is not a function of either *S*'s verbal ability or level of manifest anxiety, but is related to *Ss*' level of motivation as measured by Edwards' Need Achievement scale. The lack of relationship

between the MAS and performance is surprising in view of previously published results indicating that high MAS *Ss* show a learning decrement on complex learning tasks. However, most of the previous studies have used primarily verbal learning tasks, while this task involved a perceptual-motor type of learning. The positive findings with respect to the Need Achievement scale suggest that scores on this instrument could profitably be used as either an independent or as a control variable in future studies of perceptual-motor learning.

Summary

This study investigated the effect of the addition of high- and low-meaningful irrelevant verbal cues on the acquisition of a perceptual-motor skill. A secondary aim was to ascertain the relationship of several psychometric measures to performance on the task. It was found that the addition of verbal cues facilitated the acquisition of the skill, but the difference between the effects of high- and low-meaningful verbal cues was not significant. Performance on the task was unrelated to verbal ability or level of manifest anxiety, but was related to level of motivation as measured by Edwards' Need Achievement scale.

References

1. Battig, W. F. The effect of kinesthetic, verbal, and visual cues on the acquisition of a lever-positioning skill. *J. exp. Psychol.*, 1954, **47**, 371–380.
2. Battig, W. F. Transfer from verbal pretraining to motor performance as a function of motor task complexity. *J. exp. Psychol.*, 1956, **51**, 371–378.
3. Bendig, A. W. The development of a short form of the Manifest Anxiety Scale. *J. consult. Psychol.*, 1956, **20**, 384.
4. Dixon, W. J., & Massey, F. J. *Introduction to statistical analysis.* New York: McGraw-Hill, 1951.
5. Edwards, A. L. *Manual for the Edwards Personal Preference Schedule.* New York: Psychological Corp., 1954.
6. Kruskal, W. H., & Wallis, W. A. Use of ranks in one-criterion variance analysis. *J. Amer. Statist. Assoc.*, 1952, **47**, 583–621.
7. McAllister, D. E. The effects of various kinds of relevant verbal pretraining on subsequent motor performance. *J. exp. Psychol.*, 1953, **46**, 329–336.
8. Noble, C. E. An analysis of meaning. *Psychol. Rev.*, 1952, **59**, 421–430.
9. Warden, C. J. The relative economy of various modes of attack in the mastery of a stylus maze. *J. exp. Psychol.*, 1924, **7**, 243–275.

(Received June 25, 1956)

Journal of Experimental Psychology
Vol. 54, No. 2, 1957

OPERATOR PERFORMANCE AS A FUNCTION OF THE STATISTICAL ENCODING OF STIMULI[1]

W. D. GARVEY[2]

U. S. Naval Research Laboratory

In order to achieve maximum flow of information through the "operator channel" of a display-control system, two considerations should be given to the encoding of stimuli. The physical properties of the stimuli should be encoded to match the transmission properties of the operator channel. That is, if *S*'s "visual channel" is to be used, then stimuli should be encoded in terms of optimum brightness, hue, size, etc. Also, the statistical properties of the message source should be matched to the characteristics of the channel in such a way that those messages which convey the least amount of information may be transmitted over the channel with the greatest efficiency, and those messages which convey the most information may be transmitted over the channel with the least efficiency. A familiar example of this type of statistical encoding is found in the Morse code with its shorter code groups for frequent letters and longer code groups for the less frequent letters. The purpose of the present study is to determine the relative advantage of statistical coding during different stages of practice.

[1] The opinions or assertions contained herein are the private ones of the writer and are not to be construed as official or reflecting the views of the Navy Department or the naval services at large.

[2] The writer wishes to acknowledge the assistance of Dr. W. B. Knowles. Also he wishes to express his appreciation to Miss Jean B. Henson and Mr. E. P. Newlin, who assisted in collecting and analyzing the data.

METHOD

Apparatus.—Detailed descriptions of the apparatus have been reported (**4**). The display consisted of a 10 × 10 matrix of neon lights and the control was a 10 × 10 matrix of push buttons. The columns of lights and buttons were numbered 1 through 10 and the rows were lettered A through J. The *S*s' task was to press the control button which was identical in letter-number to the activated display light. Stimuli were automatically programmed and the mode of stimulus presentation was self-paced.

Subjects.—Thirty naval enlisted men with GCT scores between 50–55 served in the experiment.

Procedure.—The first task of *E* was to obtain a measure of what could be assumed to be an index of *S*s' "channel characteristics" when operating the particular display-control system used in this experiment. An analysis of the response times to the different stimuli (with equal stimulus probabilities) on this same system, when employed in two other experiments (**2, 3**), revealed a sharp increase in response times to stimuli going from the periphery to the center of the matrix, i.e., response times to stimuli arising in the center of the display were approximately three times as great as those in the periphery. Thus, response times to the various stimuli were taken as a measure of the transmission efficiency of the signals. As such, signals originating in the center of the display were considered to be transmitted with less efficiency over the human operator channel than those originating in the periphery. Accordingly, five different encoding procedures were established. There are 100 alternative messages, corresponding to the 100 different lights on the display. Different probabilities of occurrence were assigned to the alternative messages as follows: 36 of them were assigned probabilities so that each occurred 16 times in 760 events; each of 28 different alternatives occurred 4 times in 760 events; and the remaining 36 alternatives occurred 2 times each in 760 events.

Next, the statistical properties of this message source were matched to the transmission properties of *S*'s channel in five ways. These are shown as Codes A through E in Fig. 1. In Code A, messages with the highest probabilities were assigned to the centrally located lights on the

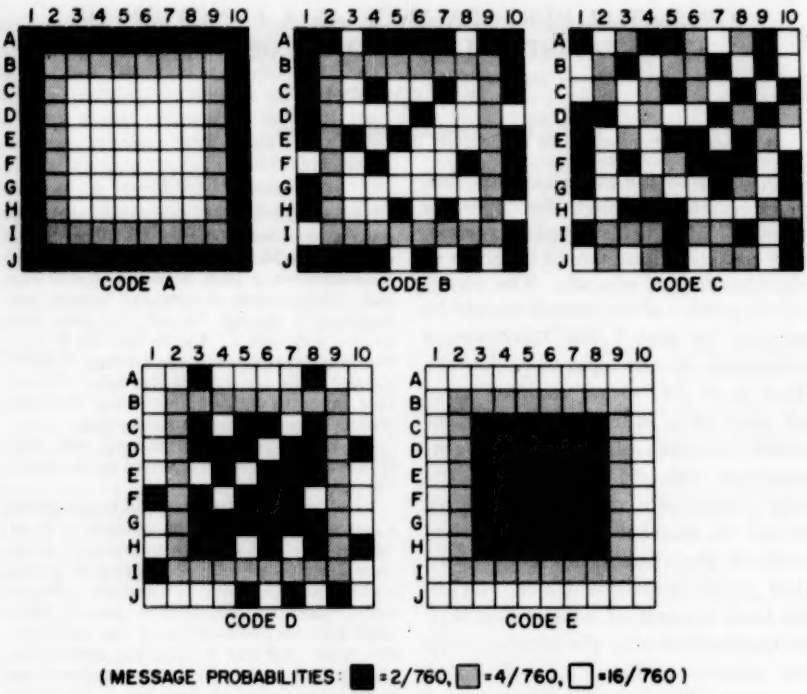

FIG. 1. Assignments of message probabilities to stimulus lights on display.

display; the messages with lowest probabilities were assigned to the peripheral lights. In Code E, the assignment was just the reverse of that of Code A, i.e., messages with the highest probabilities were assigned to the peripheral lights and those with the lowest probabilities to the central lights. In Code C, a random assignment was made. In Code B, the majority of the high probability messages were assigned to the peripheral lights, but some were assigned to the central lights as well. In Code D, the assignment was just the reverse of that of Code B, i.e., the majority of the high probability messages were assigned to the central lights. Thus, Code E is considered the optimum means of statistically encoding this message source for the characteristics of the human operator channel with this particular display-control system. The remaining codes are systematic departures from this optimum.

The *S*s were given two days' practice on a display-control system which consisted of two rows of 10 lights and two rows of 10 push buttons. The stimulus probabilities in these sessions were equal. Since previous research (**2**, **3**) indicated that performance on this system was correlated ($r = .93$) with performance on the system to be used in the experiment proper, the scores obtained during these two days were used to establish five groups of 6 *S*s which were equivalent in terms of means and variance. Each group was randomly assigned to one of the five codes.

Each *S* received all 760 stimuli each day for a period of 25 days. The stimuli were presented in random sequence and in groups of 190 per session, with 4 sessions per day. The total presentation of these 760 stimuli constituted one practice trial. The *S*s were instructed to operate the system as fast as possible without making errors. Efficiency was measured in terms of rate of responding (mean transmission time per signal).

On Day 26 the stimulus frequency was reversed for each group, i.e., lights which were assigned message probabilities of 2/760 were now

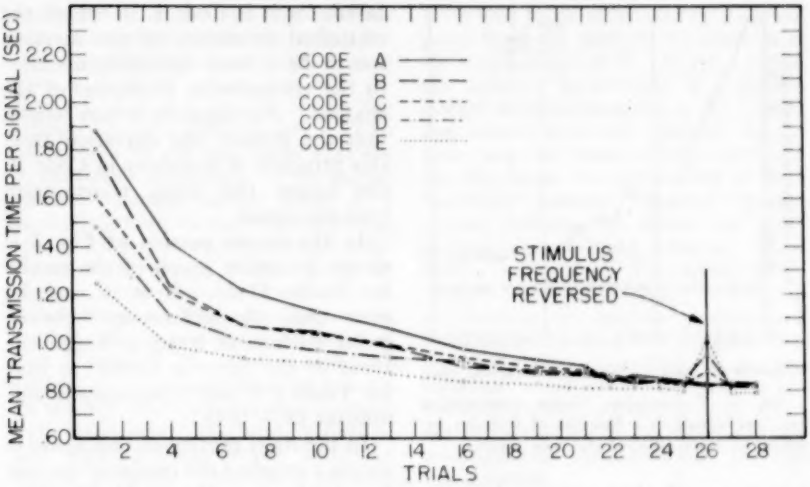

FIG. 2. Performance (mean transmission time per signal) as a function of practice with the five encoding procedures. One trial consists of the presentation of 760 stimuli.

assigned probabilities of 16/760; likewise those which were originally assigned probabilities of 16/760 were now assigned probabilities of 2/760.

On Days 27 and 28 the stimulus frequencies were reverted back to the original assignments.

RESULTS

The results are shown in Fig. 2. At the earlier stages of practice best performance was obtained with Code E where messages with the highest probabilities were matched to signals which could be transmitted with the greatest efficiency and messages with the lowest probabilities were matched to signals which could be transmitted with the least efficiency over the channel. The poorest performance was obtained with Code A, where the matching was the reverse of that of Code E. Performance for the other codes turned out as might be expected from the results with Codes A and E, i.e., Code D was poorer than Code E since some of its high probability messages were matched to some of the least efficient signals. Performance with Code C provided for average efficiency, since random matching was made between message probability and efficiency of signal transmission. And performance with Code B was better than that with Code A since some of the messages in the former code were matched to some of the most efficiently transmitted signals. In general it may be seen that there are large and systematic differences among codes at the early stages of practice, and that these differences diminish with further practice.

When performance was analyzed in terms of errors, it was found, using Wilcoxon's (**5**) nonparametric method for determining the significance of differences between two unpaired replicates, that errors were equivalent for all groups ($P > .10$). Also, errors were distributed randomly among the S-R elements and totaled less than 5% of all responses.

Figure 3 shows the effect of practice on the relationship between per-

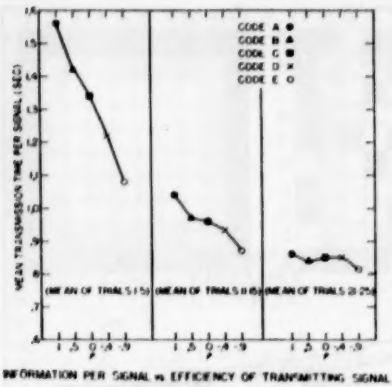

FIG. 3. Performance (mean transmission time per signal) as a function of stimulus encoding for three different stages of practice.

formance with the various codes. In order to present some scale of the relation in each code between message probability and efficiency of transmitting signals, the computed rank-difference correlations between amount of information conveyed by a message and efficiency of transmitting a signal have been plotted on the abscissa. For example, the rho of +1.0 for Code A indicates that the messages which convey the most information have been matched to the signals which are transmitted over the channel with the greatest efficiency and the messages which convey the least information to signals which are transmitted over the channel with the least efficiency.

In the left portion of the figure the data for the first five trials are shown. It may be seen that there is a systematic relationship between the five methods of coding and efficiency of performance. Using Friedman's (1) ranking test for comparison of several treatments, this function was tested and found to be significant ($P < .001$). At this early stage the desirability of statistical encoding is clearly indicated, for it is Code E in which the statistical properties of the message source have been optimally matched to the transmission properties of the channel. Furthermore, it may be seen that the greater the deviation from this principle of matching in Code E, the longer the mean transmission time per signal.

In the center portion of Fig. 2 is shown a similar graph of the results for Trials 11–15. Here it may be seen that the differences between codes are not as large, although the trend of the curve is similar to that for Trials 1–5 and is statistically significant ($P < .01$).

In the right portion of this figure is shown a graph of the results of the last five trials, i.e., Trials 21–25. The clear-cut relationship between performance with the various codes no longer exists, and the trend of this curve was not found to be significant ($P > .05$). However, using Wilcoxon's (5) unpaired replicates test, a reliable difference ($P < .05$) remains between the means of Codes A and E.

On Trial 26 the stimulus frequencies were reversed. When the differences between performance on Trials 25 and 26 were tested with matched t tests, no significant differences were observed for the two groups which had been operating the systems with Codes A and B. However, there were significant differences for those groups which had been operating the systems with Codes C ($P < .05$), D ($P < .01$) or E ($P < .01$).

DISCUSSION

During the early stages of practice there is a systematic relationship between operator performance and the extent to which the matching between message probability and signal transmission efficiency deviate from the statistical encoding of Code E. At an inter-

mediate stage of practice this relationship continues to exist; but during the final stages the relationship has broken down to such an extent that only the extremes (Codes A and E) of the five coding procedures differ reliably. These results lead to the conclusion that the "transmission properties" of the human operator channel, unlike those of a physical channel, are not fixed but change with practice. In an effort to get some notion of the nature of this change, an analysis of the response times to individual stimuli on Trials 21–25 was made for Code C. It was found, when tested with Wilcoxon's (5) unpaired replicates method, that response times did not differ reliably ($P > .05$) between frequent and infrequent peripheral elements; also, between frequent and infrequent central elements. However, response times to peripheral elements, either frequent or infrequent, were significantly shorter ($P < .05$) than the response times to the central elements. These results indicate that performance during these late stages of practice is still correlated with stimulus location, but is not correlated with stimulus frequency. However, if response times were completely independent of stimulus frequency, no change in performance should be expected by reversing the stimulus frequencies in Code C. This was not found to be the case, for it may be seen from Fig. 2 that performance deteriorated significantly for this code on Trial 26 when the stimulus frequencies were reversed. Thus, it would appear that even though response times at this late stage of practice are not correlated with stimulus frequency, the human operator channel has made some sort of adjustment to the frequencies of the individual stimulus elements through practice.

One further analysis was made of the response times to individual stimuli for each code during Trials 21–25. Using Friedman's (1) ranking test, it was found that the mean response times to the peripheral stimuli did not differ reliably ($P > .10$) among the five codes. However, the mean response times to the central stimuli showed a significant ($P < .05$) trend of increasing duration, going from the shortest for Code A to the longest for Code E. In other words, the different proportions of high frequency messages assigned to the peripheral stimuli were not reflected in the mean response times to these elements, *but* the larger the proportion of high frequency messages assigned to the central elements the shorter the mean response times to these elements. These results indicate an interaction effect, between stimulus frequency and stimulus location, on operator performance. Such an interaction effect suggests that the differential improvement in performance which took place among the five code groups may be due to the different extents to which the high-frequency messages were associated with the central elements.

The results obtained when the stimulus frequencies were reversed do not completely agree with those which should be expected if a channel with fixed characteristics were used. When the major shift was in high-frequency messages from the peripheral to the central elements, as with Codes D and E, degradation in performance would be expected and was found in the experiment. However, in the case of Code C, where the shift in frequencies from periphery to center were approximately equal, a degradation would not be expected unless some adjustment to frequency had taken place during practice. Likewise, the lack of effect with Codes A and B, where the major shift was in high-frequency messages from the central to the peripheral elements, further indicates an adjustment to stimulus frequency.

Summary

Other things equal, maximum flow of information between display and control can be obtained if the statistical properties of the message source are matched to the transmission properties of the "operator channel" in such a way that those messages which convey the least amount of information can be transmitted over the channel with the greatest efficiency and those messages which convey the largest amount of information are reserved for codes which are transmitted

over the channel with least efficiency. One of five methods of statistical encoding of stimuli used in this study adhered to this method; the remaining four methods deviated from this principle in a systematic manner.

Thirty *S*s were divided into five matched groups. Each group of *S*s was given 29 trials of 760 stimuli on a display-control system which consisted of a 10 × 10 matrix of lights and a 10 × 10 matrix of push buttons. Efficiency of performance was measured in terms of mean transmission time per signal.

The results indicated: (*a*) At the earlier stages of practice, best performance was obtained with that code where messages with the highest probabilities were matched to the most efficiently transmitted signals, and messages with low probabilities were matched to the least efficiently transmitted signals; poorest performance was obtained where the matching was the reverse of that of the most efficient code; and the other codes had intermediate efficiency. (*b*) In general, there were large and systematic differences among the code groups during the early stages of practice; but these differences diminished with practice until during the last stage of practice the systematic differences among the groups were not reliable. (*c*) The effect of reversing the stimulus frequencies was to increase the mean transmission time per signal for those code groups which were more efficient during the earlier stages of practice but to have no effect on those groups which were less efficient at the earlier stage.

References

1. Friedman, M. Use of ranks to avoid the assumption of normality implicit in analysis of variance. *J. Amer. statist. Ass.*, 1937, **32**, 675–701.
2. Garvey, W. D., & Knowles, W. B. Response time patterns associated with various display-control relationships. *J. exp. Psychol.*, 1954, **47**, 315–322.
3. Garvey, W. D., & Mitnick, L. L. Effect of additional spatial references on display-control efficiency. *J. exp. Psychol.*, 1955, **50**, 276–282.
4. Knowles, W. B., Garvey, W. D., & Baldwin, A. W. An automatic stimulus programming device. U. S. Naval Research Laboratory, Memo. Rep., 1955, No. 410.
5. Wilcoxon, F. *Some rapid approximate statistical procedures.* Stamford, Conn.: American Cyanamid Co., 1949.

(Received July 12, 1956)

Journal of Experimental Psychology
Vol. 54, No. 2, 1957

EXPECTATION AND RESISTANCE TO EXTINCTION OF A LEVER-PULLING RESPONSE AS FUNCTIONS OF PERCENTAGE OF REINFORCEMENT AND AMOUNT OF REWARD[1]

DONALD J. LEWIS AND CARL P. DUNCAN

Northwestern University

In a previous study (**2**) it was shown that there was an inverse relationship between percentage of reward and the resistance to extinction of a lever-pulling response. The present study is an attempt to extend these findings by including, along with five different percentages of reward, four different amounts of reward. In addition, three different response measures were obtained. One was, as in the previous study, the total number of plays to quitting; a second was a verbal measure of the strength of *S*'s expectancy of winning or losing on each trial; and the third was the specific motor response *S* made to operate the machine.

METHOD

Apparatus.—The apparatus was basically an electrical slot machine. It had a black crinkle-finish cabinet 22 in. long, 19½ in. high and 14½ in. deep. In the upper left-hand corner of the front surface was a slot in which quarter-sized disks could be placed. Near the top center were two lights about 3 in. apart. When lighted, the left one was green and the right one was red. In the middle center of the front surface there were 16 pushbuttons arranged in 4 columns of 4 each. The pay-off tray was at the bottom and near the center. On the left side of the box, near the bottom and front, the knob of a drawer protruded; the drawer contained the reserve supply of disks. Centered on the right side of the box was a lever, 11 in. long. The machine was played by placing a chromium-plated disk in the slot. When the *S* put a disk in the slot, the green light went on and a motor inside the box began to hum. This was the signal to push at least one button in each of the four columns of buttons. After one button had been pushed in each column, the green light went off, the red light came on, and the sound from the machine changed. The lever, which up to that point had been locked, could then be pulled. When the lever was pulled, a disk would drop into the pay-off tray (if it was a pay-off trial) and the machine was ready for the next play. As many disks as desired could be obtained from the drawer in the left side of the box; on all trials which did not pay off, disks put in the slot returned to the drawer. The *S* could not observe disks return to the drawer because the machine would not operate unless the drawer was pushed all the way in. A counter inside the machine cumulated the total number of plays.

The lights on the front of the box, the 16 buttons, and the distinctive loud sounds from inside the box actually had nothing to do with the operation of the machine. The machine could be set to pay off on any of the first 80 plays by setting a series of stops which were concealed inside the box. Thus, the machine paid off on particular plays which were determined in advance by *E*; pay-offs would not depend on anything *S* did.

Design.—There were five percentages of reinforcement and four amounts of reinforcement combined in a factorial design. Each of the 20 cells contained 10 *S*s who were given an acquisition and an extinction series. The acquisition series consisted of the first nine plays. The 0% groups received no pay-offs. The 11% group received one pay-off on the last (ninth) acquisition play. The 33% group was paid off on Plays 2, 5, and 9. The 67% group was paid on Plays 1, 2, 4, 5, 6, and 9, and the 100% group was paid on all 9 plays.

There were four amounts of reinforcement. Subjects were paid either 1¢, 10¢, 25¢, or 50¢ for each disk that they won.

The *S*s were also asked to state on each play, just before they pulled the lever, whether they expected to win or not. This was done as follows: On the upper right corner of the apparatus a strip of white cardboard was taped, on which was drawn a line 5 in. long and marked off in 1-

[1] This study was supported in part by a grant from the National Science Foundation. The authors wish to thank Dr. Arthur Thomas for his assistance.

in. intervals with the Numbers 1 through 6. Above these numbers appeared the title, "Expectation of Winning or Not Winning." Below Number 1 appeared the phrase "Fairly Sure of Not Winning," below Numbers 2 and 3 appeared "Mildly Sure of Not Winning," and "Slightly Sure of Not Winning," respectively. Below Numbers 4, 5, and 6 the phrases were the same except for the omission of "Not." On each trial *S* called out the number that corresponded to the confidence level of his expectancy. While *S* played, *E* sat to one side and recorded both the expectation *S* stated before each play and the specific buttons *S* pushed on that play.

Instructions.—The *Ss* were told that the machine was one on which they could win money if they could outguess it. They were shown the drawer containing the supply of disks and were told that, when they put one in the slot, the machine would instantly pick some combination of four buttons, one in each column, as a winner. The green light would then go on, indicating that the machine had made its pick. The *Ss* were then to press a combination of buttons, following which the red light would go on indicating that their combination had been recorded. Then they were to pull the lever, and if they had chosen correctly the machine would pay off.

The *Ss* were told how much money they would receive for each disk won and that they were to keep the disks they won separate from those they were playing with, and turn them in at the close of play for the appropriate amount of money. They were also told that they could play as long as they liked and that when they decided to quit, to turn in to *E* disks they had won and get their money. After *S* had decided to quit, he was paid off with coins, given a false but plausible "explanation" of the machine, and told not to tell any other person how the machine operated or how much he had won.

Subjects.—The *Ss* were drawn from introductory psychology courses at Northwestern University and were assigned to the various experimental groups in turn as they appeared in the laboratory. Ten *Ss* were assigned to each of the 20 experimental groups, making a total of 200 *Ss*.

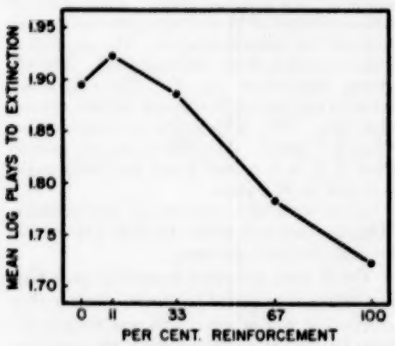

FIG. 1. The relationship between percentage of reinforcement in acquisition and the mean log plays to extinction.

RESULTS

Number of plays to extinction.—The distributions of the total number of plays to extinction were, as usual, positively skewed, but a transformation to common logarithms normalized them satisfactorily. Since a test for homogeneity of variance of the transformed data for the 20 experimental groups showed no evidence for heterogeneity, the data were treated by analysis of variance. Both percentage ($F = 2.86$, 4 and 180 *df*) and amount of reinforcement ($F = 3.32$, 3 and 180 *df*) were significant at the .05 level. Since the interaction was not significant, cells were combined and the means of points along the dimension of the two variables are shown graphically in Fig. 1 and 2. Each point in Fig. 1 is based on 40 *Ss*, in Fig. 2, on 50 *Ss*.

The *t* tests computed between the means shown in Fig. 1 indicated that the 0%, 11%, and 33% groups differed significantly (*P*'s of .02, .01, and .05, respectively) from the 100% group, and that the 11% group differed significantly from the 67% ($P = .05$). No other comparison resulted in significance. For the amount of reinforcement variable, *t* tests showed that the 1¢, 10¢, and 25¢ groups differed significantly (*P*'s of .01, .02, and .05, respectively) from the 50¢ groups but not from each other.

Acquisition expectancies.—The expectation data were treated by first combining the nine acquisition trials into three blocks of three trials each, and, because *Ss* played different

numbers of extinction plays, the extinction trials were Vincentized into 10 blocks of trials for each *S*. The expectation score for each *S* on each of the three acquisition and on the 10 extinction blocks consists of the sum of his expectation scores over that block of trials divided by the number of trials in the block. Therefore, the score for any block can be any value from 1 to 6, the same values as in the original rating scale. Any score over three represents an expectation that the machine will pay off on the next trial; a score below three represents an expectation that the machine will not pay off on the next trial. The magnitude of the score above and below three is a measure of *S*'s confidence in his expectation that the machine will or will not pay off on each successive trial.

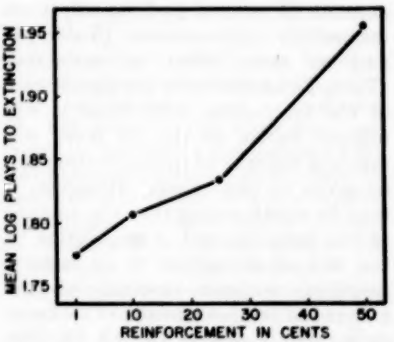

FIG. 2. The relationship between the amount of money reward and the mean log plays to extinction.

Figure 3 shows the mean expectation, both for acquisition and extinction, for the five groups varying in percentage of reinforcement. This is the main graph of expectation data and all expectation analyses are derived from it. A corresponding graph for the amount of reinforcement variable is not presented because, as will be seen, most of the statistical analyses showed little or no relation between expectation and the amount of reinforcement.

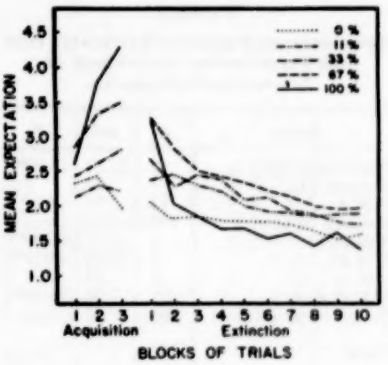

FIG. 3. The relationship between blocks of acquisition and extinction trials and mean expectation (an oral statement of confidence of winning or losing on the next trial) as a function of percentage of reinforcement. The acquisition trials are in thirds and the extinction trials Vincentized in tenths.

Because the percentage of reinforcement groups may not have started the acquisition series at the same level of expectation, an analysis of covariance was performed on the acquisition expectancies with the first block of three trials used as the control variable and the last block of three trials as the predicted variable. The *F* for amount and for the interaction between amount and percentage was less than 1. The *F* for percentage was 28.89, highly significant for 4 and 178 *df*.

A trend analysis, with the first block of three trials and the last block of three trials serving as points for the trend, gave much the same results. Table 1 summarizes this analysis. Again percentage is highly significant and amount is not. The *F* for the interaction between percentage and amount just barely achieved significance (an *F* of 1.83 is required for significance at the .05 level for 12 and 200 *df*). No clear interactional trend is apparent in the data, and because of the low level of significance obtained with a very

TABLE 1

ANALYSIS OF THE CHANGE OF EXPECTANCY FROM THE FIRST TO THE FINAL BLOCK OF ACQUISITION EXPECTANCIES

Source	*df*	*MS*	*F*
Percentage (P)	4	23.81	16.09**
Amount (A)	3	.57	
P × A	12	2.98	1.88*
Between *Ss* (error)	180	1.48	
Trials (T)	1	20.52	76.00**
T × P	4	11.05	40.93**
T × A	3	.21	
T × P × A	12	3.86	14.30**
Pooled *Ss* × T (error)	180	.27	
Total	399		

* $P < .05$.
** $P < .01$.

sensitive test, no interpretation will be attempted. Percentage of reinforcement interacted with trials in such a way, obvious in Fig. 3, that the larger percentages resulted in steeper acquisition slopes. The relationship between percentage of reinforcement and expectancy at the end of acquisition was very nearly linear, as was the relationship between percentage of reinforcement and the gain in expectancy during acquisition.

Extinction expectancies.—To determine the effect of the two major variables upon extinction expectancies a trend analysis was performed with the last block of acquisition trials, and the second and the last blocks of extinction trials making up the three points of the trend. A summary of the analysis is presented in Table 2. The significant *F* for percentage is not of particular importance since the covariance previously reported indicated that the percentage groups were significantly different at the end of acquisition. Rather, the important terms in Table 2 are the interactions between the two main variables and trials. The fact that percentage interacts significantly with trials indicates that there was a differential change in expectancies throughout extinction as a function of percentage of reinforcement. All percentage groups showed a decrease in expectancy during extinction, but the rate of decrease varied directly with the original percentage of reinforcement. This can be seen in part in Fig. 3, but appears more clearly if one plots loss in expectancy as a function of percentage of reinforcement. Two of these plots (not shown here) were made, one for loss from end of acquisition to the mean of the first two extinction blocks, the other for loss from end of acquisition to the mean of the last two extinction blocks. In both cases the losses were almost exactly linearly related to percentage of reinforcement.

Table 2 also shows that amount of reward interacts with trials. Because amount of reward had no effect on acquisition expectancies (Table 1) and no main effect on extinction (Table 2), and because the significance of the interaction with trials is significant barely at the .05 level, not much in the way of interpretation can be given to this effect. However, it may be worth noting that the amount of loss from the end of acquisition to the end of extinction is an inverse function, without reversal, of the amount of reinforcement. The mean expectancy losses for the 1¢, 5¢, 25¢, and 50¢ groups were −1.33, −1.28, −1.12, and −1.05, respectively.

Relation between the number of plays to extinction and expectancies.—Because it has been shown that the relation between percentage reinforcement and acquisition expectancies, and the relationship between percentage reinforcement and loss in expectation during extinction, are both essentially linear, one can predict that the curve relating percentage reinforcement to mean log plays to extinction (Fig. 1) will be very similar to curves relating either gain in expectation during acquisition, or loss in expectation during extinction, to mean log plays to extinction. Such relationships were, of course, obtained.

Button-pushing data.—An attempt was made to determine if the buttons pushed during acquisition and during extinction showed either any stereotypy or variability as functions of both major variables. Two response measures were used. One was the number of different buttons pressed during a block of trials; the other was the number of repetitions of buttons pressed during a block of trials. The latter measure showed no changes as a function of either major variable. For the former measure there was some evidence that the groups paid 25¢ or 50¢ showed slightly increased variability—a greater number of buttons pressed—during the first two or three extinction blocks of trials, whereas groups paid 1¢, or 5¢ showed no change.

TABLE 2

ANALYSIS OF THE CHANGE OF EXPECTANCY FROM THE END OF ACQUISITION TO THE END OF EXTINCTION

Source	*df*	*MS*	*F*
(P)	4	14.90	7.03**
(A)	3	1.90	
P × A	12	2.73	1.29
Between *Ss* (error)	180	2.12	
Trials (T)	2	73.92	246.40**
T × P	8	12.85	42.83**
T × A	6	.75	2.50*
T × P × A	24	4.97	16.57**
Pooled *Ss* × T (error)	360	.30	
Total	599		

* $P < .05$.
** $P < .01$.

DISCUSSION

Two variables were manipulated in this experiment: percentage of reinforcement and amount of reinforcement. The percentage variable was found to have a powerful effect upon both resistance to extinction and expectancies. The 0% group of this experiment confirmed the results of a previous study (**2**). No reinforcement at all results in superior resistance to extinction over much greater percentages. A discrimination hypothesis discussed in the previous study still seems adequate for these results.

The amount of reward was also found to affect resistance to extinction, but probably the major part of this effect was due to the 50¢ group. It may be that Northwestern students are unable to make sharp discriminations for money denominations below 50¢. Amount of reward had little effect upon expectancies, which is rational for this experiment, because pay-offs were never contingent upon amount, only upon percentage.

One major hypothesis that has been offered to explain the effects of partial reinforcement is the expectancy hypothesis first formulated by Humphreys (**1**). This hypothesis states that it is easier to shift from an expectancy of regular reinforcement to one of regular nonreinforcement than it is to shift from an expectancy of irregular reinforcement to one of regular nonreinforcement. Partial reinforcement entails the more difficult shift, therefore *Ss* continue longer under this condition. An analysis of the 0% group should shed some light on this hypothesis.

The *Ss* of the 0% group showed considerably more resistance to extinction than those of the 100% group. However, of the 40 *Ss* of this group, 33 maintained consistent expectancies below three throughout acquisition. Furthermore, of the 40 *Ss* in the 100% group only 18 attained a consistent expectancy of winning on the last three acquisition trials (selected because *Ss* were most apt to have attained consistency by the end of acquisition). The median number of plays to extinction of those in the 0% group who had irregular expectancies was 45, while the median number of plays of the *Ss* with regular expectancies of not winning was 71. The median number of plays to extinction of those in the 100% group who regularly expected to win during the last three acquisition trials was 57.5; the median number of plays to extinction of the 100% *Ss* who had irregular expectancies of winning during the last three acquisition plays was 49. The number of *Ss* involved in these comparisons seem too small to warrant tests of significance, but in each case the evidence indicates that the change from a trial-by-trial expectancy of regular reinforcement to regular nonreinforcement, or from an expectancy of irregular reinforcement to regular nonreinforcement does not yield the partial-reinforcement effect.

A very important qualification must be made at this point, however. The *Ss* were always asked to give their expectancies for the next trial, and the next trial only. It might still be possible that *Ss* could have such a trial-by-trial regular expectancy of not winning, and at the

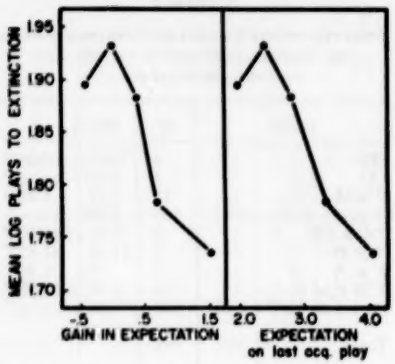

FIG. 4. The relationship between two measures of acquisition expectation and mean log plays to extinction.

same time have an *over-all* expectancy of winning sometime during the series. This means a sharp distinction must be made between trial-by-trial expectancies and over-all expectancies, and the above analysis pertains only to the trial-by-trial variety. Insofar as Humphreys' expectancies are of the over-all variety, this study has no data bearing on the problem. Even so, it has been shown that the trial-by-trial expectancies analyzed here are definitely related to percentage of reinforcement and to plays to extinction.

The relation between the two response measures—expectancies and plays to extinction—presents an interesting problem. Although one is labeled "expectancy" there should be no doubt about its being an overt response measure. This does not, of course, prevent one from inferring an underlying cognitive process if one is so inclined. For the present data, the expectancy response measure, both in acquisition and extinction, is regularly related to the independent variable, percentage of reinforcement. These relationships leave one with a curious problem. One could argue that expectancies during acquisition are not "merely" a dependent variable—a function of percentage reinforcement—but that they are also an index of an intervening variable which determines resistance to extinction. This is shown in Fig. 4, where both the gain in expectation during acquisition and the expectation value at the end of acquisition are now plotted on the baseline and mean log plays to extinction and plotted on the vertical. It can be seen that Fig. 4 is essentially the same as Fig. 1. That amount of reinforcement had an effect upon mean log plays to extinction but not upon expectancies would argue, however, against expectancies as an "intervening" variable.

SUMMARY

A total of 200 human *Ss* were assigned to 20 experimental conditions representing a factorial combination of five percentages of reward and four amounts of reward. The apparatus was an electronic slot machine which could be set to pay off according to a prearranged schedule when buttons were pushed. The percentages of reward were 0%, 11%, 33%, 67%, and 100%; amounts of reward were 1¢, 10¢, 25¢, and 50¢. Three response measures were recorded: (*a*) number of responses to extinction; (*b*) trial-by-trial expectations of winning on the next trial during acquisition and extinction; and (*c*) the button pushed on each trial.

The results showed: (*a*) The smaller the percentage of reward the more plays to extinction, except that the 0% group played slightly fewer times than the 11% group. (*b*) The larger the amount of reward the more plays to extinction. (*c*) The level of expectancies during acquisition increased linearly as a function of percentage of reward but was almost unaffected by the amount of reward. (*d*) The loss of expectancies during extinction was a linear function of the percentage of reward with the greater extinction loss following the larger percentage of reward. (*e*) The amount of reward had little effect upon extinction expectancies. (*f*) Variability of button pushing showed little effect of either percentage or amount of reward.

REFERENCES

1. HUMPHREYS, L. G. The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. *J. exp. Psychol.*, 1939, **25**, 141–158.
2. LEWIS, D. J., & DUNCAN, C. P. Effect of different percentages of money reward on extinction of a lever-pulling response. *J. exp. Psychol.*, 1956, **52**, 23–27.

(Received July 16, 1956)

Journal of Experimental Psychology
Vol. 54, No. 2, 1957

EFFECTS OF COUNTING AND ORDERING HABITS ON THE ACQUISITION OF A SIMPLE MOTOR SKILL[1]

J. H. BOWEN,[2] T. G. ANDREWS, AND SHERMAN ROSS

University of Maryland

Stimuli and responses which ordinarily occur in sequence have been used in several investigations of stimulus predifferentiation, set, and proaction. For example, Mandler (**12**) and Siipola and Israel (**18**) asked *Ss* preliminarily to learn responses to alphabetic stimuli. In experiments by Arnoult (**2**) and Gagné and Baker (**6**) letters of the alphabet were learned as responses in preliminary phases of their studies. McAllister's (**13**) *Ss* learned degree, clock, and direction responses to task stimuli. Baker and Wylie (**3**) required *Ss* to learn paired associates, one member of which specified the location and order of responses which subsequently would be learned to the stimulus specified by the other member of the pair.

Ordering or sequencing tendencies presumably are aroused when an order response is evoked or by presentation of a stimulus which is a member of a sequential set. The effects of these ordering or sequencing tendencies have not been studied either extensively or systematically. The results of some investigations, however, suggest that such habits are operative in the learning and transfer of certain discriminative-motor skills. For example, Warden (**20**) found that implicit counting of stylus maze turns was most frequently used by *Ss* as an aid in maze learning. Rexroad's (**15**) *Ss* counted choice sequences during the learning of a multiple-choice test. The present investigation attempted, within the general framework of cue-producing response theory (**14**), to study the effects of the preliminary arousal of counting and ordering habits on the subsequent acquisition of a simple discriminative-motor skill.

The experiment followed the paradigm of stimulus differentiation. After a preliminary training period with the independent variables specified below, all *Ss* learned a task which involved a sequential congruency between the display stimuli and the control responses. The independent variables were: (*a*) preliminary application of spatially sequential and nonsequential verbal labels to task stimuli, (*b*) three levels of preliminary label learning, (*c*) two types of stimulus labeling, letters and numerals, (*d*) preliminary instructions concerning the sequential nature of the display-control relationships, and (*e*) three levels of preliminary exposure to task stimuli (stimulus exposure warm-up).

METHOD

Subjects.—The *Ss* were 180 enlisted men who had been assigned to the U. S. Army Ordnance School, Aberdeen Proving Grounds, Aberdeen, Maryland. The age range of the sample was from 17 to 32 yr., with a median age of 20 yr. The *Ss* were obtained through an accidental sampling process which was carried out by noncommissioned officers.

Apparatus.—The apparatus consisted of a 2 × 2 ft. display panel, on which were mounted eight 1-in. jewel lights, and a control panel with eight toggle switches. The spatial arrangement of stimuli and of response switches is shown in

[1] This experiment is one of a series on performance decrement supported in part by Contract No. DA-49-007-MD-222 between the Medical Research and Development Division, Office of the Surgeon General, Department of the Army, and the University of Maryland.

[2] Now at the Naval Research Laboratory, Washington, D. C.

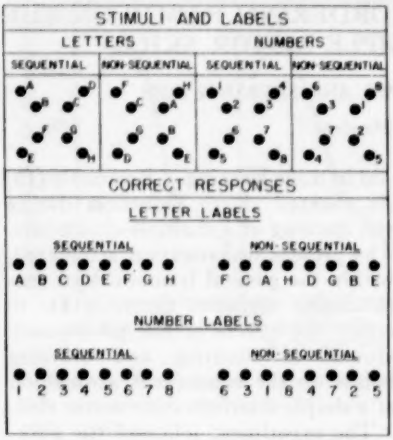

FIG. 1. Schematic of stimulus-response relationships indicating stimulus labels and correct responses.

Fig. 1. The display was placed 4 ft. from *S* and the control was placed on a table in front of him. The letters and numerals beside each black circle in the upper half of Fig. 1 represent the labels which were learned for respective stimuli. The letters and numerals below each black circle in the lower portion of Fig. 1 indicate the response switch which would extinguish the stimulus having the corresponding letter or numeral label. Neither the display nor control was marked with these letters or numerals.

General procedure.—The procedure was divided into two phases for all experimental groups. During the preliminary phase, or predifferentiation training (PDT), some *Ss* learned letter or numeral labels for display stimuli. The labels had either a sequential or nonsequential characteristic and were learned to three levels of proficiency. The *Ss* in the labeling groups were subdivided further as to whether they did or did not receive special instructions. These instructions were designed to produce an understanding of display-control relationships. There were two types of control groups. One set of controls was given three levels of preliminary stimulus warm-up (warm-up controls), both with and without special apparatus instructions. The remainder of the controls received no preliminary stimulus experience (the no-exposure controls) or were given only the special instructions.

The second phase of the experiment, which produced the dependent variable data and was administered to all *Ss*, required that *S* learn the correct switch which would extinguish each display stimulus. This phase followed the preliminary phase described above and was called the motor learning (ML) task. Each of the experimental, control, and ML procedures is considered below.

Verbal predifferentiation.—In verbal-label PDT, *S* was seated before the display and was instructed that he was to learn the label for each of the display stimuli. The *E* then illuminated the stimuli singly for a 3-sec. interval, with 3 sec. between stimuli. If *S* knew the letter or numeral label for a stimulus he verbalized that response as soon as possible after stimulus presentation. Regardless of *S*'s response, *E* verbalized the correct label for the stimulus at the end of the presentation interval. The same relationships prevailed between labels within the sequential letter and numeral groups and within the nonsequential letter and numeral groups. For example, the letter "H" is the eighth letter of the alphabet and therefore was learned as the label for the same stimulus as the numeral "8." Label sequence refers to the *spatially* sequential characteristic of the labels, as learned for the stimuli, and does not refer to the order in which the stimuli were illuminated by *E*.

The display stimuli were illuminated in an irregular order which was predetermined by *E*. A practice trial was defined by the illumination of each of the eight stimuli. Each stimulus appeared once within a trial and no stimulus was presented twice in immediate succession. The *Ss* in verbal-labeling PDT groups received 2, 7, or 12 practice trials in learning the labels for their group. These three levels of learning were chosen on the basis of pilot study results. On a learning function the three levels represent, respectively, slight learning, the inflection point, and approximately 70% overlearning of the labels.

Verbal instructions, which were designed to produce an understanding of display-control relationships, were also administered by *E* as a PDT procedure.[3] The *S* was shown a simplified wiring diagram of the apparatus and *E* concomitantly stressed the following points: (*a*) there was a simple series relationship in the circuit between each display stimulus, each of *S*'s control switches, and each of *E*'s programming switches, and (*b*) if *S* pressed the correct switch, the series circuit would be interrupted and the stimulus would be extinguished. As *E* gave these instructions he traced out the

[3] Copies of the complete doctoral dissertation from which this article is adapted, containing instructional materials and stimulus programming details, may be obtained from University Microfilms, Ann Arbor, Michigan. Remit $1.41 for microfilm and request Publication No. 14,181.

electrical pathways which were involved on the circuit. For the groups which received both the labeling PDT as well as the instructional PDT, the instructional PDT followed verbal-label learning and preceded the ML phase.

Stimulus warm-up controls.—Each *S* in the stimulus warm-up control groups was seated before the display panel and was instructed that: (*a*) the stimuli would be lighted, one at a time, (*b*) *S* should watch the lights closely, and (*c*) *S* should try to determine whether there was some relationship between the stimuli. The *S* watched the stimuli flash on, remain lighted for 3-sec., and go off. There was a 3-sec. interstimulus interval. The *S* was instructed not to verbalize overtly during this procedure. The stimuli were lighted in the same irregular presentation order which was used for verbal-label PDT. There were three levels of stimulus exposure warm-up which corresponded to the same number of stimulus presentations received by *S*s in groups given 2, 7, and 12 trials of verbal-label PDT. When the special display-control instructions were administered in addition to the stimulus warm-up, these instructions followed the warm-up and immediately preceded ML.

No-exposure controls.—There were two no-exposure control groups. One learned the ML task without prior exposure to display stimuli, stimulus labeling, or special display-control instructions. The other group learned the ML task following the special display-control instructions, but without any additional prior experience.

Motor learning task.—The ML task required that the *S* learn the location of the switch on the control panel which would extinguish an illuminated stimulus. To avoid premature attentive and postural adjustments directed toward the control panel, all *S*s began their experimental experience with the control panel covered by a wooden cover. In addition, the *S*s in experimental groups receiving PDT or warm-up were not informed of any relationships between their prior experience and the ML task. As in the PDT and warm-up conditions, stimuli were presented for 3-sec., with a 3-sec. interval between presentations. The *S* could press only one switch for each stimulus presentation, and *E* did not inform *S* of the correct switch in case of error. The *S*s in all groups received 15 ML trials, with a trial defined by presentation of each of the eight stimuli. The *E* recorded the number of errors made by each *S* in each trial. Figure 1 shows that, regardless of the type of prior experience, the four switches on the left half of the control panel controlled the four stimuli in the upper half of the display, while the four switches on the right half of the control panel controlled the four stimuli in the lower half of the display. Thus, it may be seen that the sequential PDT labels were designed to produce an ordered congruency relationship between the stimuli and the subsequently learned ML responses. On the other hand, the nonsequential labels were expected to be disruptive of any congruency between stimuli and the ML responses.

Design.—The experimental variables were combined in a $3 \times 2 \times 2 \times 2$ factorial design. The respective variables were: (*a*) three levels of label learning, 2, 7, and 12 trials; (*b*) two degrees of sequence in labels, sequential (S) and nonsequential (NS); (*c*) two types of labels, letters and numerals; and (*d*) two levels of display-control relationship instructions, instructions and no instructions. Of 180 *S*s, 120 were assigned randomly to the cells of this design, with 5 *S*s for each treatment combination. The remaining 60 *S*s were apportioned randomly to two independent designs. One design was produced by combination of the two levels of display-control instructions with 2, 7, and 12 trials of stimulus exposure warm-up, and each of these six combinations received 5 *S*s. The second design divided 30 *S*s into two groups, one which learned the ML task following display-control instructions, and one which received no prior experience of any type.

Results

The dependent variable in this study is the number of incorrect responses made within each trial of ML. Principal interest centers upon error data which were obtained during the first ML trial. Performance on this trial is believed most clearly to reflect the presence or absence of PDT and warm-up treatment effects. An estimate of persistent PDT effects is obtained through an examination of error data from intermediate ML trials. There were heterogeneity of within-groups variances, correlations between means and variances, and skewness in the experimental group data for ML Trial 1. These data were normalized and were subjected to variance analysis. The results are presented in Table 1.

The following conclusions are supported by the analyses in Table 1. First, the sequential character of the

TABLE 1

ANALYSES OF VARIANCE OF NORMALIZED ERROR DATA FOR TRIAL 1 OF MOTOR LEARNING AND OF UNTRANSFORMED POOLED ERROR DATA FOR TRIALS 6–10 OF MOTOR LEARNING

Source	*df*	Trial 1 *MS*	Trials 6–10 *MS*
Level of PDT (L)	2	.26	9.78*
Label Sequence (S)	1	154.81**	93.63**
Type of Label (T)	1	.09	10.32
Instructions (I)	1	4.92	7.91
L × S	2	14.27**	2.34
L × T	2	.22	1.69
L × I	2	3.67	1.49
S × T	1	14.09*	27.27**
S × I	1	.02	14.41*
T × I	1	.01	.05
L × S × T	2	5.19	3.65
L × S × I	2	2.99	3.91
S × T × I	1	8.14	4.76
L × T × I	2	8.19	3.76
L × S × T × I	2	27.89**	1.44
Within Groups (Error)	96	2.47	2.69
Total	119		

* Significant at the .05 level.
** Significant at the .01 level.

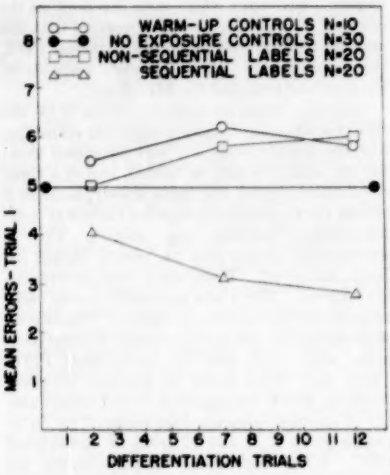
WARM-UP CONTROLS N=10
NO EXPOSURE CONTROLS N=30
NON-SEQUENTIAL LABELS N=20
SEQUENTIAL LABELS N=20
MEAN ERRORS - TRIAL 1
DIFFERENTIATION TRIALS
1 2 3 4 5 6 7 8 9 10 11 12

FIG. 2. Mean errors on Trial 1 of discriminative-motor learning as a function of number of differentiation trials for experimental and control groups.

PDT verbal labels was a significant variance source in learning new responses to the labeled stimuli. Second, the effects of label sequence were dependent upon the level of PDT. Examination of Trial 1 error means suggests that ML error was an increasing function of the level of nonsequential PDT, while ML error was a decreasing function of the level of sequential PDT. These relationships are depicted in Fig. 2. Third, the effects of label sequence were dependent upon the type of label employed. In this instance, error means showed that there was a greater difference between the effects of PDT with sequential and nonsequential numeral labels than between the effects produced by sequential and nonsequential letter labels. The use of nonsequential numeral labels produced more ML error than nonsequential letter labels, and sequential numeral labels lowered ML error more than sequential letter labels. In all instances in which variance analysis revealed the significance of a variable, examination of error means showed them to be in the expected directions.

The statistical significance of the differences between the performance of the PDT groups and the performance of the no-exposure and warm-up controls was determined. No significant differences were found between display-control instructional subgroups, between letter-numeral subgroups, and between the warm-up subgroups. Data for these groups were combined in subsequent statistical analyses for both Trial 1 and Trial 6 data. Warm-up data are not combined for graphic purposes in Fig. 2. When it was impossible to accept the null hypothesis concerning pairs of variances, the differences between error means for the experimental and control groups were tested

TABLE 2

RESULTS OF *t* TESTS BETWEEN EXPERIMENTAL AND CONTROL GROUP ERROR MEANS FOR TRIALS 1 AND 6 OF MOTOR LEARNING
(Display-control instruction, warm-up level and letter-numeral subgroup data are combined)

	Warm-up Control		2-S		2-N-S		7-S		7-N-S		12-S		12-N-S	
	1	6	1	6	1	6	1	6	1	6	1	6	1	6
No-Exposure Control	—	—	—	—	—	—	**	**	—	—	**	**	*	—
Warm-up Control			**	—	*	—	**	**	—	—	**	**	—	—
2-S					—	—	—	**	**	—	—	**	**	—
2-N-S							—	**	—	—	**	**	—	—
7-S									**	**	—	**	**	**
7-N-S											**	**	—	—
12-S													**	**

Note.—Levels of PDT are denoted as 2, 7, and 12. The letters S and N-S refer, respectively, to sequential and nonsequential label groups.
* Significant at the .05 level.
** Significant at the .01 level.

with a statistic developed by Cochran and Cox (**4**, p. 92). Table 2 presents the results of the *t*-tests between experimental and control group means.

The following conclusions concerning Trial 1 data are supported by the tests in Table 2. First, with the exception of two trials of sequential PDT, all levels of sequential PDT produced significant positive transfer effects in subsequent ML. Second, the positive transfer effects of sequential PDT were maximized by seven PDT trials, i.e., there were no statistically significant differences between the ML effects produced by 7 and 12 PDT trials. Third, only 12 trials of PDT with nonsequential labels produced negative transfer effects in ML.

The error data from ML Trials 6–10 were combined to stabilize means and error variances and were subjected to variance analysis. In addition, the error data for Trial 6 of ML were examined with *t* tests as were data for Trial 1 of ML. Table 1 presents the results of the variance analyses.

The conclusions to be drawn from this section of Table 1 are that, in later ML: (*a*) as in Trial 1 data, the sequential characteristic of the PDT labels was a significant source of variability; (*b*) as in Trial 1 data, the effects of label sequence were dependent upon the type of label employed; (*c*) the level of PDT became a significant variance source; and (*d*) the effects of label sequence became dependent upon the presence or absence of display-control instructions. In the case of the latter conclusion, examination of error means revealed that the differences between the effects of sequential and nonsequential labeling were smaller in conjunction with display-control instructions than were these differences in the absence of instructions. Thus, exposure to display-control instructions following nonsequential PDT reduced the superiority of sequential over nonsequential PDT effects on ML.

Table 2 also presents the results of *t* tests between experimental and control group means for data from Trial 6 of ML. The conclusions supported by these analyses are that, on Trial 6 of ML: (*a*) as in Trial 1 data, all levels of sequential PDT except two

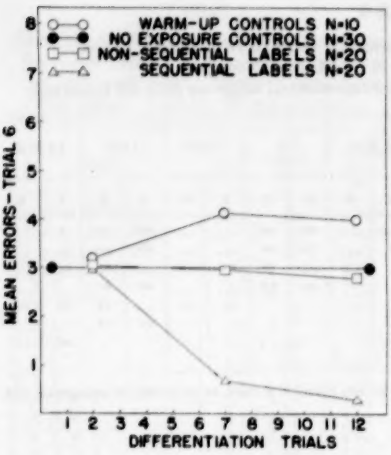
WARM-UP CONTROLS N=10
NO EXPOSURE CONTROLS N=30
NON-SEQUENTIAL LABELS N=20
SEQUENTIAL LABELS N=20
MEAN ERRORS - TRIAL 6
DIFFERENTIATION TRIALS
1 2 3 4 5 6 7 8 9 10 11 12

FIG. 3. Mean errors on Trial 6 of discriminative-motor learning as a function of number of differentiation trials for experimental and control groups.

trials produced significant positive ML transfer effects; (*b*) the positive transfer effects of sequential PDT were maximized by 12 PDT trials, i.e., there was a significant difference between the ML effects of 7 and 12 trials of sequential PDT; and (*c*) the negative ML transfer effects of 12 trials of nonsequential PDT had disappeared. Figure 3 depicts the relationships between experimental and control group means for Trial 6 data. As in Fig. 2, warm-up data are not combined in Fig. 3 but letter-numeral and display-control instructional subgroup data are combined.

DISCUSSION

The results of this study are interpreted to be in support of the assumptions of cue-producing response theory (**14**) and stimulus differentiation theory (**8**). It is assumed that, because of their ready availability and strong cue-producing properties, counting and ordering habits contributed a high degree of distinctiveness to the task training stimuli. In addition, these habits are assumed to have served as a generalized mediational vehicle when new responses were learned to old stimuli. Thus, the arousal of counting habits is believed to have evoked components of motor sequences which were useful when the new reactions were learned. A comment by Miller and Dollard (**14**, p. 76) is illustrative in this connection:

> ". . . the Roman numeral III, the Arabic numeral 3, and the written word three are innately quite distinctive as cues. But training in our society has attached the same cue-producing responses to each of these patterns of stimulation. The most obvious of these cue-producing responses is the verbal one; but there are a host of other responses, such as holding up three fingers, which tend to occur in minimal form, supply additional cues common to these three stimuli, and mediate generalization from any one to the others."

The greater superiority of sequential over nonsequential labels in conjunction with numerals leads to the inference that those habits which are concerned with numeral responses are of greater strength than the habits mediating letter responses. These results support the deduction that the components which are related within a given habit sequence are organized into habit families with regard to interfamilial strength.

It should be noted that observing response theory, as formulated by Spence (**19**), does not accurately predict certain results of this investigation. Spence has emphasized that the observing response is transferred as a general receptor orientation and, in this, his view is similar to Irion's (**10**) principle concerning transfer of nonspecific set. In the present study, general receptor orientations were clearly aroused by the stimulus warm-up trials, and more specific postural, attentive, and other adjustments were aroused by display-control instructions. However, prior stimulus warm-up produced no effects in subsequent learning and the display-control instructions were only effective in combination with prior stimulus labeling.

The finding with respect to warm-up and instructions is in accord with results obtained by Ammons (1), working with the pursuit rotor, and by Hamilton and Mola (9), working with a maze. They found no significant warm-up effects. On the other hand, it has been shown (e.g., 11) that appropriate warm-up activities can significantly reduce retention loss for nonsense syllables. There seems no ready way of compromising these discrepant results. It is probable that the threshold for the arousal of facilitative postural and 'attentive adjustments varies widely between tasks and depends, not only on S-R similarity relationships, but also upon task complexity, difficulty, verbal constitution, and the initial discriminability of task stimuli.

A number of investigations have stressed the importance of the *spatial* correspondence of stimuli and responses in the production of S-R congruency or compatibility (e.g., 5). The present study has shown that an additional factor—the *S*'s repertoire of counting and ordering habits—is experimentally manipulable such that S-R congruency is produced. This conclusion is given support by an incidental finding by Garvey and Knowles (7). During the learning of a display-control ensemble, their *S*s often ignored the coding which was made available as an integral part of the ensemble. The *S*s substituted their own coding devices. These investigators suggested that *S*s' codings be classified as *perceptual-analytic*, *spatial-verbal*, and *iconic-symbolic*. The *S*s in the present experiment probably used codings which could be classified in the *spatial-verbal* category. The *S*s in another prediffer-entiation study (17), when questioned, revealed that they had been looking for and identifying portions of the stimuli during ML. Hence, their techniques might be classified as *perceptual-analytic*.

In the present study when *S*s learned nonsequential PDT labels to task stimuli, a slight, significant, and transient negative ML effect resulted. This finding agrees with those of other investigators. One group of *S*s in McAllister's (13) experiment learned nonsense syllable labels for task stimuli (the relevant stimulus group) to which new reactions subsequently were learned. The ML performance of this group contained a significantly large number of errors. Also, when Mandler's (12) *S*s learned a new reaction to familiar stimuli, the result was a negative transfer effect: Several concepts have been used in the past to describe the fact that performance of a task can be inhibited by preliminary training with the task stimuli. They have included McAllister's (13) concept of the relevance of preliminary training, Rees and Israel's (16) notion of task intrinsic sets, the concept of stimulus differentiation as proposed by Gibson (8), and the population stereotype concept (21). These formulations emphasize that when the habits which are preliminarily aroused are appropriate to task performance, then the period required in searching for and in fixating a new reaction is reduced. Thus, in this study when nonsequential labels had been applied, and inappropriate habits presumably aroused, then there was an error increase in subsequent learning.

Summary

This study investigated the effects on the acquisition of a switch-pressing response of: preliminary application of sequential and nonsequential labels to task stimuli, three levels of preliminary label learning, two types of stimulus codings, two levels of preliminary instructions concerning display-control relationships, and three amounts of stimulus exposure warm-up. One control group learned the switch-pressing task without prior stimulus exposure, instructions, or labeling experience. There was a sequential display-control relationship between task stimuli and responses in the common motor task. The dependent variable was the number of errors within each trial of motor learning. The results, interpreted within the framework of cue-producing response and stimulus differentiation theory, are:

1. The preliminary coding of task stimuli with sequential labels produced a stable error reduction in subsequent acquisition of the common motor task. In both early and late acquisition stages the effects of label sequence were dependent upon the type of label. Numerals

were more effective in error reduction than letters.

2. In early and late response acquisition stages, error reduction appeared to be a decreasing negatively accelerated function of the level of preliminary stimulus labeling. In early acquisition the effects of sequential labeling were maximized by 7 labeling trials. In late acquisition the effects were maximized by 12 labeling trials. Twelve labeling trials with nonsequential labels produced a slight but significant error increase in early acquisition trials, but in later trials there were no effects which were due to nonsequential labeling.

3. There was an interaction of display-control instructions and the sequential characteristic of the labels. Exposure to the instructions following nonsequential labeling reduced the superiority of sequential over nonsequential label effects.

4. Display-control instructions, stimulus exposure warm-up, and letter and numeral coding did not, considered independently, produce effects when the new responses were learned to task stimuli.

References

1. Ammons, R. B. Effects of pre-practice activities on rotary pursuit performance. *J. exp. Psychol.*, 1951, **41**, 187–191.
2. Arnoult, M. D. Transfer of predifferentiation training in simple and multiple shape discrimination. *J. exp. Psychol.*, 1953, **45**, 401–409.
3. Baker, K. E., & Wylie, R. C. Transfer of verbal training to a motor task. *J. exp. Psychol.*, 1950, **40**, 632–638.
4. Cochran, W. G., & Cox, G. M. *Experimental designs.* New York. Wiley, 1950.
5. Fitts, P. M., & Seeger, C. M. S-R compatibility: spatial characteristics of stimulus-response codes. *J. exp. Psychol.*, 1953, **46**, 199–210.
6. Gagné, R. M., & Baker, K. E. Stimulus predifferentiation as a factor in transfer of training. *J. exp. Psychol.*, 1950, **40**, 439–451.
7. Garvey, W. D., & Knowles, W. B. Response time patterns associated with various display-control relationships. *J. exp. Psychol.*, 1954, **47**, 315–322.
8. Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.
9. Hamilton, C. E., & Mola, W. R. Warm-up effect in human maze learning. *J. exp. Psychol.*, 1953, **45**, 437–441.
10. Irion, A. L. The relation of "set" to retention. *Psychol. Rev.*, 1948, **47**, 336–341.
11. Irion, A. L., & Wham, D. S. Recovery from retention loss as a function of amount of pre-recall warming-up. *J. exp. Psychol.*, 1951, **41**, 242–246.
12. Mandler, G. Transfer of training as a function of degree of response overlearning. *J. exp. Psychol.*, 1954, **47**, 411–417.
13. McAllister, D. E. The effects of various kinds of relevant verbal pretraining on subsequent motor performance. *J. exp. Psychol.*, 1953, **46**, 329–336.
14. Miller, N. E., & Dollard, J. *Social learning and imitation.* New Haven: Yale Univer. Press, 1941.
15. Rexroad, C. N. Verbalizations in multiple choice reactions. *Psychol. Rev.*, 1926, **33**, 451–458.
16. Rees, H. J., & Israel, H. E. An investigation of the establishment and operation of mental sets. *Psychol. Monogr.*, 1935, **46**, No. 6 (Whole No. 210).
17. Rossman, I. L., & Goss, A. E. The acquired distinctiveness of cues: the role of discriminative verbal responses in facilitating the acquisition of discriminative-motor responses. *J. exp. Psychol.*, 1951, **42**, 173–182.
18. Siipola, E. M., & Israel, H. E. Habit interference as dependent upon stage of training. *Amer. J. Psychol.*, 1933, **45**, 205–227.
19. Spence, K. W. Continuous versus non-continuous interpretations of discrimination learning. *Psychol. Rev.*, 1940, **47**, 271–288.
20. Warden, C. J. The relative economy of various modes of attack in the mastery of a stylus maze. *J. exp. Psychol.*, 1924, **7**, 243–275.
21. Warrick, M. J. Direction of motor preferences in positioning visual indicators by means of control knobs. *Amer. Psychologist*, 1947, **2**, 345. (Abstract)

(Received July 24, 1956)

Journal of Experimental Psychology
Vol. 54, No. 2, 1957

CONTINUOUS PERSPECTIVE TRANSFORMATIONS AND THE PERCEPTION OF RIGID MOTION [1]

JAMES J. GIBSON AND ELEANOR J. GIBSON

Cornell University

An exploratory survey of the varieties of optical motion which could serve as stimuli for the perception of motions in the world (**6, 7**) suggested the hypothesis that one kind of geometrical motion in a plane yields an impression of a rigid motion in space but that any other kind of geometrical motion in a plane does not. The stimulus pattern was always a "texture," that is, a grouping of dark shapes on a light background. If, on a motion picture screen, it underwent a *continuous sequence of perspective transformations* in any of six ways, it gave a perception of a rigid surface moving in one of six ways—the three transpositions and three rotations which, in combination, exhaust the possibilities of mechanical movement. If it underwent continuous transformations *not* of this geometrical kind (but only a few examples were presented) it aroused perceptions of nonrigid or elastic surface motions, of the kind exemplified in the movements of organisms.

Of the six rigid phenomenal motions, three (rotation around the line of sight, transposition up or down, and transposition right or left) are induced by a stimulus which common sense would call motion; one (transposition along the line of sight) by a stimulus which common sense would call expansion or contraction; and only the other two (rotation around a horizontal or a vertical axis) by a stimulus which common sense would call a *transformation*. Optics, however, demands geometrical terms. All six projected motions are different parameters of continuous perspective transformation, and they are mathematically akin. Common sense tells us that the first three optical motions *should* give the perceptions they do (a motion yields a motion) and that the last three should not (how can a change of size or shape yield a motion?). The assumption is that a visual experience has to resemble visually the optical stimulus that produced it. But a better assumption is that experiences need only correlate with their stimuli, not replicate them, and the present hypothesis says that any continuous sequence of perspective transformations is the correlate of perceptually rigid motion.

There was in the film some evidence to suggest that this hypothesis must be qualified if the perspective transformations are those obtained with parallel projection instead of polar projection, that is, with the special case of transformations when the focus of projection is at infinity. The two apparent rotations around a horizontal or a vertical axis then seemed to become somewhat ambiguous as to rigidity or elasticity, and apparent reversals of direction

[1] This experiment was reported by the first author as part of an address entitled *Stimulation and Perception* delivered as retiring President of the Division of Experimental Psychology, APA, in September, 1955. The work was supported in part by the Office of Naval Research under Contract NONR 401(14) with Cornell University. An early form of the apparatus to be described was constructed, and preliminary experiments were performed by H. R. Cort. The writers are also obligated to Dr. O. W. Smith for ideas and assistance.

of rotation appeared. The apparent approach or recession also fails of necessity in this case because the change of size of the stimulus disappears with parallel projection. If the above observations are verified, the hypothesis should specify perspective transformations with polar projection. Previous experimental work on the kinetic depth effect (**14, 15**) or on other appearances of depth in moving fields (**3, 12**) does not supply evidence for or against the amended hypothesis since in general the changes of shape studied in them were not polar projective. These experiments, moreover, are mainly concerned with what can be called the appearance of *internal depth* of an object, whereas what we are here talking about is the appearance of *slant depth* of the face of an object. The distinction is made clear in the film. The *apparent motion in depth* previously studied by Smith (**13**), however, is relevant to our hypothesis. One or two of Wallach's many experiments on the kinetic depth effect (**14,** p. 212 ff.) are relevant indirectly if the changes of such impoverished stimuli as line segments or angles are restated in terms of perspective transformations.

This hypothesis is comparable to, but different from, the principle involved in Wertheimer's "law of common fate" (**16**) in several respects. Both refer to some kind of motion in or of a grouping of spots or forms, but Wertheimer's law predicts the organization of a figure in the visual field, whereas this predicts the quality of rigidity of a surface or surface-like experience in space. Wertheimer's law seems to imply that the various parts of the complex are united by sharing a common motion (such as moving in the same direction with the same velocity) but this hypothesis asserts that any perspective transformation is a single motion mathematically, including the size and slant transformations where, analytically considered, every part moves in relation to every other. Wertheimer's law leads to experiments on "configurations of motions" (**9, 12**) in which each part of the complex undergoes components of translation or rotation but the part is not itself transformed; a geometrical transformation, however, is something that permeates every part as well as the whole of a texture, and the apparatus used in the present experiment satisfies this condition.

It might be noted that the problem of how we discriminate the rigidity of rotating solid objects and of approaching or receding solid objects in the environment is closely connected with the traditional problems of shape constancy and size constancy. Langdon has recently shown that the shape constancy of an object is considerably increased under conditions highly unfavorable for it when the object is made to rotate (**10**). Likewise the question of why we see a rigid environment when we move among the solid surfaces around us is closely connected with the traditional problem of space perception (**8**).

Aim of the experiment.—The experiment to be reported sought to answer four questions. First, does the appropriate parameter of continuous perspective transformations with polar projection always give the perception of the changing slant of a constant shape? Second, are the judgments of amount of change of slant away from the picture-plane in good psychophysical correspondence with the "extent" or "length" of the transformation sequence? Also, how variable are these judgments? Third,

is the outcome dependent on or independent of the *kind* of shape or texture on which the transformation is imposed? Fourth, how accurate, if at all, is the judgment of slant away from the picture-plane when only the static end product of the transformation sequence is presented to *O* but not the motion leading to it?

Method

Apparatus and Stimuli

The optical geometry of the apparatus used is shown in Fig. 1. The device can be termed a "shadow transformer." Essentially, it presents to an eye an optic array of limited scope within the boundaries of which either static patterns or continuous perspective transformations can occur. In this optic array, unlike those of everyday vision, the differential light intensities and their structure are under *E*'s control; the pattern is the same for either eye, and the need for differential convergence and accommodation is eliminated. All the "cues for depth," in short, tend to determine a flat plane except those of form and motion, which are thus isolated for study. The source of this converging array is a window in a translucent screen.

This optical stimulus is artificially produced by the diverging ray sheaf from a point source of light, into which shadows are introduced by opacities of one sort or another attached to a transparent plane mount. Rotations or translations of the mount (on bearings or tracks outside the ray sheaf) yield corresponding transformation sequences of the shadow. This experiment utilized rotation on a vertical axis. The stimuli were the mirror reversals of these moving shadows, visible on the other side of the translucent screen. If an apparent rotation of a "virtual object" is induced by such a stimulus, it should always be opposite to the rotation of the shadow caster, without ambiguity.

The seated *O*, in a dimly illuminated room facing a large white surface, saw a luminous square window 36 cm. on a side at a distance of 180 cm., made of translucent plastic ⅛ in. thick. The light source was fixed at the same distance behind the window as the eye was in front. It was a 300-w. Sylvania point source carbon arc lamp, but any lamp with a single filament of small diameter (up to 1 mm. or more) will serve the purpose. The window was visibly flat. Binocular vision was permitted *O* after preliminary work failed to show any difference between the use of one or two eyes. The mounts were transparent rectangular sheets of ¼-in. plastic, of such size (30 × 100 cm.) that when they were centered and rotated on a turntable placed midway between the point

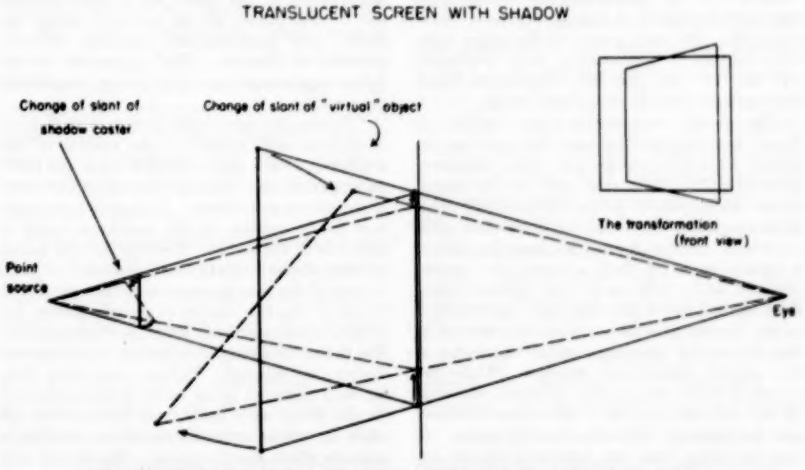
TRANSLUCENT SCREEN WITH SHADOW
Change of slant of shadow caster
Change of slant of "virtual" object
The transformation (front view)
Point source
Eye
VIEW OF APPARATUS FROM ABOVE PRODUCING A SLANT TRANSFORMATION

Fig. 1. The shadow transformer.

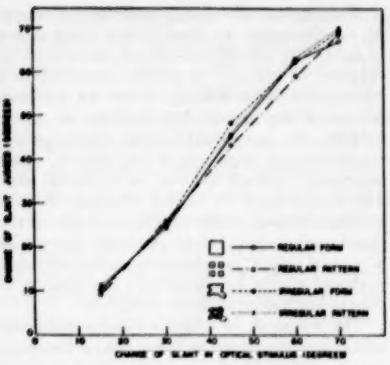

FIG. 2. Judgments of change of slant as a function of the length of the transformation sequence.

source and the window they could be turned 70° from the parallel plane without the edges being projected within the window. The turntable could be rotated back and forth through an arc of variable length by an adjustable eccentric linkage, geared to a motor with a variable speed drive. A speed which gave 2-sec. cycles of semirotation was chosen, after exploration indicated that an optimum might be in this neighborhood, although the rate was not critical for the experiment. The quantitative variable of this experiment, then, was the "length" of the transformation sequence, as expressed in degrees of angular excursion of the turntable. We shall return to this point later. Five degrees of semirotation were presented: 15°, 30°, 45°, 60°, and 70°. Each cycle began with and returned to the parallel plane.

The forms transformed.—The variety of forms, patterns, and textures that can be projected with this device has been suggested elsewhere (**7**). Four were used in the experiment: an amoeboid group of amoeboid dark shapes or spots (the irregular texture), a solid amoeboid contour form (the irregular form), a square group of dark squares (the regular texture), and a solid square (the regular form). Each was cut out of gummed paper and attached to the central area of a transparent mount so that its shadow was projected to the center of the square translucent window. With the mount parallel, the regular shadows extended 20 cm. each way in the 36 cm. square window, and the irregular shadows about the same. It may be noted that the "regular" stimuli are constituted of rectilinear contours and alignments and the "irregular" stimuli of randomly curved contours and alignments. There are also differences in symmetry, and perhaps other geometrical properties. The "forms" are bounded by a single closed contour and the "textures" by many closed contours; the total contour length is much greater in the latter stimuli. A texture might be described as a "form of forms," as distinguished from a form as such. These textures were, however, very "coarse"; there were 36 squares in the "platoon" and 36 "amoebas" in the "colony."

The variable protractor.—For recording judgments of change of slant, *O* had before him a sort of protractor with its baseline parallel to the plane of the screen. It bore an adjustable pointer which could be moved to indicate an angle of semirotation. The top side was blank but the bottom side carried another pointer and a scale which could be read accurately by *E* after each trial.

Instructions and Procedure

The experimental group.—Before receiving any formal instructions for the experiment, each *O* was seated and told: "You see in front of you a screen with a window in it which will be illuminated during the experiment. I will first show you a moving pattern of dark lines filling the window. If what you see is a movement of some kind of object, describe it."

A network (woven wire fencing of a common type) was then placed on the turntable and turned through various excursions. Although the question was intended to suggest neither a deformation in the plane nor a rigid rotation out of the plane, all *O*s reported seeing the latter, and spontaneously reported different amounts of rotation. The suggestions in the following instructions were hence considered permissible:

"During the experiment proper, a dark form or pattern will appear in the middle of the window. It will seem to rotate back and forth on a vertical axis—to turn away from the plane of the screen and return. Your task is to judge how far it rotates, or the maximum angle it makes with the screen. Use the circular model in front of you to make this judgment."

One of the four patterns was then presented at one of the five degrees of transformation for 20 sec., which permitted 10 cycles of stimulation. The *O* had no difficulty in making his judgment during that interval. Twenty such trials (five for each pattern) in an order counterbalanced for the group were made, and then another 20 trials in reverse order to determine whether a practice effect would appear. The *O* was not told his errors. There were 20 *O*s in the group.

The control group.—A separate group of 30 *O*s was treated as similarly as possible except

that the four stimulus patterns were motionless. Only the end product of each transformation sequence was presented, and only one degree of transformation was used—that with the mount at 60°. For the preliminary exposure, *O* was shown a motionless pattern filling the window, half the group seeing the network of lines and the other half a less objective cloud-like pattern (this making no difference in the outcome) and he was asked if he saw an object of some sort. Then *O* was told that he would see a form or pattern in the middle of the window. It might be parallel or slanted away from the screen. If he saw it slanted away from the plane he was asked to judge the angle it made using the model in front of him. Four trials were given (one for each pattern) in an order counterbalanced for the group.

Results

The experimental group.—The first question is whether all *O*s saw the changing slant of a rigid shape. As stated above, all did at the outset. During the 40 trials which followed, each of considerable duration, many spontaneous descriptions were offered, and 8 of the 20 *O*s observed at some stage that the display could be seen as a compression of a two-dimensional pattern. They were all psychologists. Twelve did not so report, and stated at the end that they had never observed it. The two-dimensional impressions did not persist long enough to prevent the requested judgments of changing slant. There was no difference in this respect between the regular or irregular forms or textures.

The second question is whether the judgments of change of slant are a function of the amount of change of form. The "length" of the transformation sequence is expressed as the inverse angular excursion of the shadow caster, and this variable is plotted on the horizontal axis of Fig. 2. The judgments are plotted on the vertical axis, each point representing a mean of 40 reproductions. The

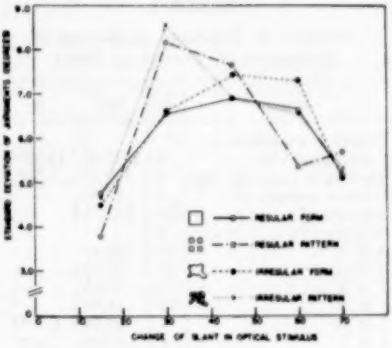

Fig. 3. Variability of judgments as a function of the length of the transformation sequence.

function is linear, except for a small tendency to underestimate (about 4 to 5°) the 15° and 30° angles. Even more striking, however, is the similarity of the functions for the regular and irregular and also for the single and multiple stimuli. This suggests that a transformation can be responded to as such, independently of what gets transformed.

A related question is how variable the judgments are within the group. Figure 3 shows the *SD*'s of the judgments as a function of the change of slant "in" the stimulus. The variable errors range between about 4° and 8°. They rise, but fall again as the maximum slant begins to approach 90°, at which limit a phenomenal surface becomes an edge.

The graph does not show any obvious differences between the errors obtained with the four kinds of pattern used. An analysis of variance (Table 1), however, is necessary to discover whether either multiplicity or regularity of the patterns is significantly related to errors of judgment. Multiplicity or "texturedness" is not. The effect of regularity

TABLE 1

ANALYSIS OF VARIANCE OF ERRORS OF JUDGMENTS OF CHANGE OF SLANT

Source	*df*	*MS*	*F*
Degree of excursion of turntable (A)	4	1437.07	13.12**
Regular or irregular (R)	1	320.68	5.56*
Form or texture (T)	1	72.90	
Subjects (*S*)	19	453.93	
Practice (P)	1	.53	
A × *S*	76	109.47	
R × *S*	19	57.69	
T × *S*	19	52.63	
P × *S*	19	23.07	
A × R	4	60.54	4.77**
A × T	4	67.09	2.66*
T × R	1	13.15	
A × *S* × R	76	12.69	
A × T × *S*	76	25.24	
T × R × *S*	19	299.13	

Note.—Only those interactions which appeared to be of some interest or were used for the *F* tests (**11**, p. 330 and 338) are included in the Table.
* P = .05.
** P = .01.

is significant at only the 5% level.[2] The tendency is in the expected direction but is weak, considering predictions that might be made from Gestalt theory about "good form." It is strengthened, perhaps, by the significant interaction between angle and regularity which seems to reflect the tendency, barely noticeable in the graphs, for the irregular forms to depart slightly more from linearity at the larger angles. All the forms in this experiment were apparently good enough to carry the transformation, and it was this which mainly determined the judgments. This answers the third question. The form of the *change* seems to be what is important, not the form itself. A conception of the various forms that optical change may take is, however unfamiliar, probably necessary for an understanding of the perceptual process.

[2] The 5% level is probably not acceptable here, since some inhomogeneity of variance is evident in the data.

It may be noted from Table 1 that no significant practice effect appeared between the first and second blocks of 20 trials. They have been pooled in Fig. 2 and 3. The two halves of the data independently warrant the same conclusions, when the curve of Fig. 2 is plotted separately for them.

The control group.—The outcome of the control experiment was radically different inasmuch as the judgments of slant depended on the regularity of the form or texture presented. The irregular stimuli, in fact, generally appeared in the plane of the screen (85% of 60 judgments) while the regular stimuli generally appeared at a slant from the screen (97% of 60 judgments). Even for the regular stimuli, however, the mean degree of slant perceived was only 24° (*SD* about 12°) whereas for the moving regular stimuli the mean had been 61° (*SD* about 6°). This is gross underestimation for the motionless and great accuracy for the moving stimuli. The underestimation of slant is consistent with previous research on static optical forms and optical textures under similar conditions. A trapezoidal form can sometimes arouse an impression of slant, but an exact linkage between the apparent shape and the apparent slant (a "psychological invariant") is not obtained (**2**). A static optical texture with a compression of texture on one meridian relative to the other induces a perception of surface slant, but even when the texture is regular the slant is underestimated, and when the texture is less regular the slant is more underestimated (**4**, p. 380).

The irregular form and the irregular texture displayed in this experiment were evidently not of such a kind as to appear slanted when altered by a slant transformation, since they generally still looked frontal. The *O*s,

TABLE 2

MEANS AND *SD*'S OF JUDGMENTS OF CHANGE OF SLANT FOR TWO *O*'S

Angular Excursion of Turntable	Solid Square				Square of Squares				Solid Amoeba				Group of Amoebas			
	O_1		O_2		O_1		O_2		O_1		O_2		O_1		O_2	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
15°	8.4	3.0	9.4	3.3	9.1	2.4	7.9	2.2	9.7	4.6	9.2	1.7	8.7	2.6	9.3	4.1
30°	20.7	4.9	25.0	7.1	21.3	6.2	26.2	4.4	21.3	4.8	27.9	9.0	22.0	4.6	26.3	6.2
45°	47.3	8.0	43.2	7.1	39.8	7.6	44.1	4.1	45.5	5.4	44.4	5.2	41.8	9.5	43.2	5.9
60°	66.2	5.4	59.5	6.1	64.2	4.8	63.6	4.9	66.6	5.0	58.0	5.6	66.1	7.7	58.4	5.2
70°	69.2	5.6	60.8	4.6	72.2	3.7	67.5	9.1	76.6	3.4	65.2	6.0	78.2	2.6	68.2	6.7

Note.—Each cell in the table represents 10 judgments.

of course, had never seen them *before* transformation.[3] A truer statement of the matter is that the family of perspective transformations of the amoeboid stimuli has no unique member with immediately identifiable properties. The family of perspective transformations of the quadrilateral stimuli *does* have such a member—the square. At the outset there is one in this family which the other members can be transformations *of*, but none in the former family. Hence, rectilinear contour and alignment—a rectilinear structure—provide a primary basis for presumptions of slant which nonrectilinear structure does not.

One of the writers has argued that there exists a better basis than contour for presumptions or even perceptions of surface slant, namely, internal texture and its density variables (**4**).[4] It might be expected, therefore, that the texture of amoeboid spots would induce slant more often than the amoeboid form, and the texture of squares more slant than the square. Neither expectation was fulfilled. The explanation may be that the textures used (36 elements, or only about 6 each way) were too coarse to make the density variables determinate.

Evidently a configuration with only the feeblest stimulating power for depth perception, or none at all, can nevertheless carry a transformation sequence which yields accurate depth perception. This answers the fourth question.

Individual results.—Time did not permit more than two judgments for each condition per *O* in the experiment just reported. It was thought desirable to run a more exhaustive series in order to determine the extent of the variable error of the slant judgments for single individuals. Accordingly, two *O*'s were run for five days each, so that 10 judgments per condition were available.[5] Table 2

[3] An *O* who had become familiar with the frontal aspect of the irregular form or texture (such as either *E*) could see another aspect as slanted. So also could an *O* who had persistently observed these patterns undergoing continuous transformations. Presumably these *O*'s had learned to *identify the surface*, i.e., to recognize a previously unfamiliar object in nonfrontal aspects. This, we believe, is Wallach's "memory effect" in the perception of tridimensional forms (**15**).

[4] In the study referred to, the correlate of slant was said loosely to be a "gradient" of texture density. The correlate might as well be expressed as an unequal density along two meridians, or the ratio of these densities, or a special sort of compression of texture. These are all comprehended in the geometrical notion of a perspective transformation.

[5] Mr. John Hay kindly obtained these data.

shows the means and *SD*'s of these judgments. The means are strikingly like those for the group, especially for O_1, who underestimates the smaller angles and is more accurate or overestimates on the larger. The second *O* tends to underestimate throughout. All the curves, when plotted, are close to linearity, and the particular pattern makes little difference. The *SD*'s are of the same order as for the group.

Discussion

Kinetic depth effect and memory.— Wallach believes that the kinetic depth effect of what we call a transformation sequence must be due to some effect of past experience on present momentary experience (**15**, p. 364). He argues that since any single member of the sequence looks flat in isolation, the present member has depth only because the memory traces of past members enter into the present perception. He assumes that only the present member of the sequence can be a stimulus for the eye. This is a perfectly logical extension of the classical theory which strictly separates "traces" and "stimuli" as determiners of perception (or behavior). But does it not reduce the theory to an absurdity? Does a stimulus last for a second, a millisecond, or a microsecond? And what about the doctrine that "a stimulus is always a change"? Is it not theoretically preferable to suppose that a transformation is a stimulus in its own right, just as a nontransformation is a stimulus? Or, still better, that sequence, as well as pattern, is a variable of stimulation? Is it not better to take the risk that traces might vanish from psychological theory than the risk that the stimulus might vanish? Perhaps the addiction of Gestalt theorists to the concept of traces is what has prevented some of them from studying temporal forms as effectively as they have spatial forms.

Wallach has convincingly demonstrated (and the present observations confirm it) that a motionless nonsense pattern of a certain sort on a shadow screen which at the outset usually looks flat will usually look deep after *O* has seen it in transformation. This result may well be considered a genuine effect of memory, or at least of recognition. But it is far from proving that memory traces are necessary for the perception of rigid motion. It might prove not an effect of memory on depth perception but an effect of memory on the depth interpretation of an ambiguous static picture. It is possible that the role of *learning* in space perception is quite a different affair, consisting not of the enriching of bare stimuli but the discriminating of rich stimuli.

The anchoring of a transformation sequence and the identifying of the object.— A motionless pattern of sufficient irregularity appears in the plane of the screen even when its shadow-making pattern is slanted to the screen. What, then, would happen to the judgment of *change* of slant if the shadow caster were rotated between, say, 30° and 60° instead of between 0° and 60°? Would *O* then see the pattern as departing from and returning to the frontal parallel plane? Is the neat correspondence between length of transformation sequence and change of slant shown in Fig. 2 destroyed in these circumstances? These questions lead into fundamental problems of psychophysical scaling and shape constancy; complete answers cannot be given, but enough evidence has been obtained to be worth reporting here.

If the 30° to 60° cycle is presented to a naive *O* with a rectilinear or regular pattern, judgments are made with some accuracy. The apparent surface does not begin and end in the parallel plane. If the cycle is presented with an irregular pattern, however, the naive *O* reports a large change of slant which seems difficult to estimate, and the object *does* seem to return to the parallel plane. The crucial experience seems to be his first view of the motionless pattern. He sees an object ***but his first impression of it is an object thinner than the shadow-casting pattern.*** All observations so far suggest

that if he is thus led to identify at the outset an object of the "wrong" shape, the judgments of change of slant will persistently be out of scale. His scale of slant will be stretched and displaced, as it were, until he is given the opportunity of anchoring it at 0° and 90° which, at the same time, enables him to identify the rigid shape of the object.

The size and distance of the virtual object are *never* determined by the optical stimulus in our experiment. Similarly, the slant and shape of the virtual object may be misdetermined at the outset by showing it, even briefly, as a continuous nontransformation—a static frontal pattern. This establishes a false shape constancy for the phenomenal object. When it rotates, anomalies of space perception will occur of the sort demonstrated strikingly by Ames with the "rotating trapezoidal window" (**1**). If care is taken by *E*, however, to avoid the procedure above, the evidence indicates that an intermediate transformation cycle can be correctly judged from the outset, whenever the length of the transformation sequence is sufficient. If a naive *O* is first shown an anchored transformation sequence from 0° to 90°, or even if he is first shown the motionless pattern at 45° but is told not to assume that the object is necessarily in the plane of the picture, then a 30° to 60° cycle is judged approximately as such. The tentative conclusion is that a motion consisting of a perspective transformation sequence can determine both a definite rigid shape and a definite change of slant in perception, for a wholly unfamiliar object, without the need of any presumption whatever about the probable shape of the object based on memory (**2**).

It sounds very paradoxical to assert that a change of form of the stimulus can yield a constant form with change of slant in the percept. The paradox probably arises because two different meanings of the word "form" are employed (**5**), the first being geometrical and the second substantial. In any event, the assertion is inaccurate since the evidence now suggests that the *form* of the change of form of the stimulus is what determines the perception of rigid motion.

If the impression of surface rigidity in visual perception can be accounted for, the constancy of shape of objects is explained at the same time. The face of a unique solid object is ordinarily given not as a form but as a unique family of transformations to the eye. The difference between one solid object and another, in contrast, is *not* given as a relation of perspective transformation, nor is the difference between an earlier and a later state of a physically *changing* object. The perceptual problem of the recognizing or identifying of unchanging objects by their shape has to be approached in the light of these facts.

Summary

Continuous perspective transformations of varying length were presented in 2-sec. cycles to each *O* on the visibly flat surface of a translucent screen. Judgments of the amount of change of slant of the apparently rigid object were in good correspondence with the length of the transformation sequence, without depending on the kind of pattern which carried the transformation. The patterns differed with respect to regularity vs. irregularity and form vs. texture. Regularity may have had a small effect on the variability of judgments but texturedness did not.

As a control, the same patterns were presented motionless at the end of a transformation sequence. In general the irregular patterns appeared to be in the frontal plane but altered in shape; the regular pattern appeared at some degree of slant, but the judgments were not accurate. Evidently impressions of changing slant are precise whereas corresponding impressions of unchanging slant are ambiguous or weak. Rectilinear contours and alignments seem to provide some basis for impressions of unchanging slant. A sequence of perspective transformations, on the other hand, seems to yield an impression of changing slant whether or not such regularity is present.

For irregular unfamiliar patterns there was evidence to suggest that the perceiving of the rigid shape of the virtual object is intimately connected with the perceiving of change of slant of the object. Misidentification of the shape at the outset was accompanied by anomalies in the perception of slant.

The eye appears to be very sensitive to a continuous perspective transformation in the optic array. Psychophysical experiments are possible if the parameters of this stimulus are isolated and controlled.

References

1. Ames, A., Jr. Visual perception and the rotating trapezoidal window. *Psychol. Monogr.*, 1951, **65**, No. 7 (Whole No. 324).
2. Beck, J., & Gibson, J. J. The relation of apparent shape to apparent slant in the perception of objects. *J. exp. Psychol.*, 1955, **50**, 125–133.
3. Fisichelli, V. R. Effect of rotational axis and dimensional variations on the re-reversals of apparent movement in Lissajous figures. *Amer. J. Psychol.*, 1946, **59**, 669–675.
4. Gibson, J. J. The perception of visual surfaces. *Amer. J. Psychol.*, 1950, **63**, 367–384.
5. Gibson, J. J. What is a form? *Psychol. Rev.*, 1951, **58**, 403–412.
6. Gibson, J. J. *Optical motions and transformations as stimuli for visual perception.* (Motion picture film). State College, Pa.: Psychol. Cinema Register, 1955.
7. Gibson, J. J. Optical motions and transformations as stimuli for visual perception. *Psychol. Rev.*, in press.
8. Gibson, J. J., Olum, P., & Rosenblatt, F. Parallax and perspective during aircraft landings. *Amer. J. Psychol.*, 1955, **68**, 372–385.
9. Johannson, G. *Configurations in event perception.* Uppsala: Almqvist and Wiksell, 1950.
10. Langdon, J. The perception of a changing shape. *Quart. J. exp. Psychol.*, 1951, **3**, 157–165.
11. McNemar, Q. *Psychological statistics.* (2nd ed.) New York: Wiley, 1955.
12. Metzger, W. Tiefenerscheinungen in optischen Bewegungsfeldern. *Psychol. Forsch.*, 1935, **20**, 195–260.
13. Smith, W. A. Sensitivity to apparent movement in depth as a function of "property of movement." *J. exp. Psychol.*, 1951, **42**, 143–152.
14. Wallach, H., & O'Connell, D. N. The kinetic depth effect. *J. exp. Psychol.*, 1953, **45**, 205–217.
15. Wallach, H., O'Connell, D. N., & Neisser, U. The memory effect of visual perception of three-dimensional form. *J. exp. Psychol.*, 1953, **45**, 360–368.
16. Wertheimer, M. Untersuchungen zur Lehre von der Gestalt. II. *Psychol. Forsch.*, 1923, **4**, 301–350.

(Received July 27, 1956)

Journal of Experimental Psychology
Vol. 54, No. 2, 1957

STUDIES OF DISTRIBUTED PRACTICE: XVI. SOME EVIDENCE ON THE NATURE OF THE INHIBITION INVOLVED IN MASSED LEARNING OF VERBAL MATERIALS

BENTON J. UNDERWOOD[1]

Northwestern University[2]

Previous studies (**4**, **5**, **7**) have shown that when distributed practice facilitates the learning of serial nonsense lists and when interlist interference is relatively low, massed practice is followed by better retention after 24 hr. than is distributed practice. This fact and a number of related facts still lack theoretical solution. The present study was designed to reduce the number of possible alternative solutions.

One of the most appealing and widely used theoretical notions is that of an inhibition which builds up during work and dissipates rapidly during rest. In this general form the notion has been used to explain certain facts of motor learning, conditioning, and rote learning, although the particular mechanism by which the inhibition is generated has varied. When applied to rote serial learning (e.g., **2**) the construct has been used to account for facilitation of learning by distributed practice and for reminiscence. When *S* learns under relatively massed practice it is assumed that an inhibition develops which, while not retarding the development of associative strength, depresses performance. Therefore, performance is poorer under massed practice than under distributed practice since the inhibition is assumed to dissipate rapidly during the short distribution intervals. If a short rest interval is introduced during massed practice, the inhibition will likewise presumably dissipate and performance will increase on the trial after rest (reminiscence). Finally, following performance at the same criterial level by massed and by distributed practice recall of the massed list should be superior to that of the distributed list.

As mentioned above it has been found that with serial nonsense or consonant lists distributed practice facilitates acquisition. Better 24-hr. retention following massed learning than following distributed learning has also been found when interlist interference is low. These two facts are obviously in accord with the inhibition notion. The third phenomenon, reminiscence, has not been investigated in the experimental situations which have produced these other two phenomena. The purpose of the present study was to make a test for reminiscence in such situations. The reasoning is that if reminiscence is found in this situation the use of the simple inhibitory construct in an explanatory system would be indicated; if reminiscence is not found it would seem necessary to change the characteristics ascribed an inhibition or omit it from the system.

Concerning reminiscence as such, it is well known that many investigators have failed to obtain the phenomenon even with materials which result

[1]Rudolph Schulz has supervised the collection of the data and both he and Jack Richardson have critically read the manuscript.

[2]This work was done under Contract N7onr-45008, Project NR 154-057, between Northwestern University and the Office of Naval Research.

in differences in acquisition by massed and distributed practice. The critical difference between such studies (e.g., **1**) and those which have shown reminiscence (e.g., **8**) has not been determined. The present study may be considered as one more attempt to delimit the situations in which reminiscence does and does not occur. The experimental situation used is one (**7**) in which wide differences in acquisition rate are produced by massed and distributed practice and in which massed practice is followed by better retention than is distributed practice.

Method

General.—The reference study (**7**) is one in which serial consonant lists were used. The massed-practice group of 84 *Ss* from this study will be used as the reminiscence control group in the present report. This control group learned List 1 to one perfect recitation on Day 1 and recalled and relearned this list after 24 hr. On Day 2 they learned two additional lists, on Day 3 three lists, and on Day 4 a single list (List 7) which was recalled after 24 hr. Since learning of all lists was by massed practice, the performance of these *Ss* will serve as a control for the present experimental groups where a single short rest interval is introduced at a specified point in learning. If reminiscence occurs, performance of the experimental groups should be better than the control on the trial immediately after the rest interval. Reminiscence intervals have been introduced for the experimental groups on the first list and on the seventh list. Since the lists have high interlist similarity, change in performance over a short rest interval can be examined on the first list where interlist interference is low and on the seventh where it is high. For each critical list an interval is introduced early in learning and late in learning to determine if reminiscence is a function of degree of learning preceding the interval.

Specific.—The serial lists consisted of six consonant syllables, but since the first item served as an anticipatory cue, only five items were actually learned. These lists have been presented elsewhere (**7**). They were presented at a 2-sec. rate on a Hull-type drum with a 2-sec. interval between trials defining massing. Learning was by the anticipation method. Each *S* learned seven lists; the lists were rotated so that each occurred equally often at each of the seven positions in the series.

There were two experimental groups. In one group a 2-min. rest interval was introduced after *S* achieved a criterion of two correct responses on a single trial (2/5). This interval was introduced during the learning of Lists 1 and 7. A second group had the 2-min. interval introduced after a criterion of four correct responses on a single trial was achieved (4/5). The interval for this group was also introduced during the learning of Lists 1 and 7 only. For both groups Lists 2–6 constituted only filler lists to build up interlist interference. The 2-min. reminiscence interval was used because when reminiscence has been found it is at a maximum at 2 min. Except for the reminiscence intervals the two experimental groups were treated in exactly the same way as the control group. Following the rest interval, learning was continued to one perfect trial and recall and relearning of Lists 1 and 7 took place after 24 hr. There were 28 *Ss* in each experimental group and, as will be subsequently shown, the learning ability of these two groups did not differ appreciably from that of the 84 *Ss* in the control group.

During the 2-min. reminiscence intervals *S* was given a symbol cancellation task (**3**). In order to insure a smooth transition from learning to symbol cancellation and from cancellation to learning, *S* was given practice in starting and stopping the cancellation before learning the first list. The cues for starting and stopping were the same as those used for the reminiscence interval. The *S* was told that he would have such a symbol-cancellation period during the learning of List 1 and was told the same before the learning of List 7.

Results

Learning.—The basic learning measures are shown in Table 1 in terms of mean number of trials required to reach three criteria, namely, 2/5, 4/5, and 5/5 on a single trial. These measures are shown for both Lists 1 and 7. For each criterion for each list F's were determined for the three groups. None of these even approached significance, the largest being 1.42. For most of these sets of three distributions, variance was significantly heterogeneous although on all subsequent F's and t's to be reported this was not true. The ex-

TABLE 1

MEAN NUMBER OF TRIALS TO REACH THREE CRITERIA IN LEARNING LISTS 1 AND 7

Group	Criteria					
	2/5		4/5		5/5	
	Mean	σ_M	Mean	σ_M	Mean	σ_M
List 1						
2/5	7.18	1.23	21.18	2.74	31.25	2.53
4/5	5.89	.64	18.32	1.72	27.39	2.22
Control	6.63	.69	22.45	1.68	33.19	2.13
List 7						
2/5	2.86	.43	11.39	1.20	18.11	1.39
4/5	2.86	.72	13.25	1.60	18.68	1.89
Control	3.37	.32	13.64	.85	21.20	1.13

perimental group learning to the 4/5 criterion seems to be largely responsible for this heterogeneity in learning as shown in Table 1. However, it does not seem likely that this denies the two conclusions to be drawn from these data. First, the three groups did not differ significantly in learning ability as judged by the number of trials required to reach a criterion of two correct responses on a single trial. Second, the introduction of a single 2-min. rest interval has no appreciable influence on rate of learning to a subsequent criterion. Thus, the introduction of a 2-min. interval after reaching two correct responses did not influence the rate of learning to four correct responses, and the introduction of the interval after four correct did not influence the rate of learning to one perfect. In each case the direction of the differences is in favor of the experimental groups but, as mentioned earlier, none of the differences is significant. Separate statistical tests of each experimental group against the control did not change the conclusion.

Reminiscence.—The most specific test of reminiscence can be made by comparing preformance of the experimental group with the control group on the trial immediately following the 2-min. interval. Thus, performance of the two groups is compared at exactly the same point, the only difference in treatment being the rest for the experimental group just prior to this point. To do this, the number correct on the trial immediately following the rest was subtracted from the number correct on the trial immediately preceding rest. This was done for each *S* separately and a distribution of gain scores derived, or as is true for the present data, loss scores. Corresponding scores were obtained for the control *S*s; that is, the difference in performance on the trial when *S* attained a criterion of two correct responses (and four correct) and the immediately following trial was determined. If the loss scores are less for

TABLE 2

LOSS SCORES FOR THE TWO EXPERIMENTAL GROUPS AND THE CONTROL GROUP AT TWO CRITERIA FOR LISTS 1 AND 7

Condition	List 1		List 7	
	Mean	σ_M	Mean	σ_M
Exp.—2/5	.68	.16	.82	.21
Control—2/5	.54	.11	.33	.12
Exp.—4/5	2.11	.26	1.71	.27
Control—4/5	1.39	.15	1.30	.14

the experimental group than for the control, it would be considered evidence for reminiscence. In determining the difference scores for the control group, one problem was encountered. Some *Ss* jumped from a criterion of three correct responses to five correct (one perfect). For these *Ss*, therefore, no "postrest" trial was available. To handle this for such *Ss* the "prerest" performance score was considered five correct and the "postrest" score 2.5 correct, the latter being the average of *Ss* who had not jumped the criterion for four correct. For List 1 there were 5 *Ss* who had jumped, for List 7, 10. It can be seen that the method of handling these cases would, if any bias is introduced, tend toward enhancing the likelihood of obtaining reminiscence.

The mean loss scores (difference scores) for each group for each criterion for Lists 1 and 7 are shown in Table 2. In every case the loss for the experimental groups is greater than for the control. In the case of the criterion of four correct on List 1 the difference in loss between the experimental and control groups gives a t of 2.48, indicating significant forgetting. Determinations have also been made of the number of *Ss* showing a gain (rather than a loss). In every comparison the percentage was greater for the control group than for the corresponding experimental group. Examination of the data in terms of serial position shows not the slightest evidence of reminiscence at any serial position. The evidence seems conclusive that no reminiscence occurred.

Recall.—All *Ss* recalled Lists 1 and 7 after 24 hr. These results will be mentioned briefly. In conformance with the previous study (7), recall of List 1 was much higher than the recall of List 7. The three groups did not differ significantly on recall of either List 1 or List 7. However, on the recall of List 1 the control group recalled more than the experimental groups and on List 7 this was reversed. To determine if this interaction was significant, the recall for each *S* on List 7 was subtracted from his recall on List 1. This was done for each group and the resulting three distributions showed an F of 5.60. With 2 and 137 df this is beyond the 1% significance level. The same "switchover" was evident in the relearning scores with the F significant at the 5% level. These findings are mentioned since they indicate that a single rest period produces the same effect as does fully distributed practice as shown in the previous study (7). The magnitude of the effect, however, is appreciably less.

DISCUSSION

No evidence for reminiscence has been found in these data. For this particular study, the implications lie not in why no reminiscence was found, but in their relevance for a theory which attempts to account for the effects of distributed practice. In a situation in which marked acceleration in acquisition rate is produced by distributed practice and in which massed practice gives better retention than distributed practice (when interlist interference is low), the failure to find reminiscence may be used to delimit the nature of an inhibitory con-

struct which might be postulated. The following points summarize some of the possible conclusions concerning an inhibitory construct.

1. If massed practice develops inhibition which depresses performance, it does not seem possible to postulate a rapidly dissipating function for the inhibition.

2. It would be possible to postulate an inhibition which occurs with massed practice which actually retards growth of associative strength and which dissipates rapidly with rest. Better recall of a massed list than a distributed list would then be handled by other mechanisms or by allowing the inhibition to produce some performance decrement in addition to retarding the growth of associative strength. Actually, the better recall of massed lists over distributed lists is not great and seems to be mostly confined to strong items learned by slow-learning *Ss* (**4, 6, 7**).

3. Another alternative is to consider the possibility that distributed practice inhibits retention when interlist interference is low. There is no conclusive way to tell whether massed practice facilitates retention or distributed practice inhibits it; all that is known is that there is a difference. Thus, it may be that distributed practice inhibits recall of strong items learned by fast *Ss*.

There are, of course, other ways of approaching the problems. The main point to be made by the present study is that it does not now seem justifiable to postulate a rapidly dissipating inhibition which depresses performance and does not influence associative strength. It should be remembered also that when interlist interference is high, retention is better following distributed practice than following massed (**7**); this strongly suggests that other mechanisms will have to be considered in any complete explanatory system.

Summary

This study was an effort to delimit the nature of the inhibition which might be postulated as an explanatory mechanism for the facts associated with massing and distribution of verbal materials. Studies which show faster acquisition by distributed practice and better recall following massed practice would suggest a simple inhibitory construct which depresses performance and dissipates rapidly with rest. If these characteristics of such a construct are adequate, reminiscence over a single short rest interval should occur during massed practice. Therefore, reminiscence was studied with serial consonant syllables, five items per list, the reminiscence interval being 2 min. and introduced early in learning and late in learning for different groups. These intervals were introduced in one case where interlist interference was low, and in the other when it was high. A control group with no rest intervals served as the base for determining whether or not reminiscence occurred.

The results show no reminiscence; indeed, forgetting was observed over all 2-min. rest intervals although in only one case was this forgetting statistically significant. It is concluded that if a performance-depressing inhibition is postulated to account for poorer performance under massed than under distributed practice, this inhibition cannot be assigned rapidly dissipating properties.

References

1. Archer, E. J. Retention of serial nonsense syllables as a function of rest-interval responding rate and meaningfulness. *J. exp. Psychol.*, 1953, **45**, 245–252.
2. Hull, C. L., Hovland, C. I., Ross, R. T., Hall, M., Perkins, D. T., & Fitch, F. B. *Mathematico-deductive theory of rote learning.* New Haven: Yale Univer. Press, 1940.
3. Underwood, B. J. Studies of distributed practice: VI. The influence of rest-interval activity in serial learning. *J. exp. Psychol.*, 1952, **43**, 329–340.
4. Underwood, B. J. Studies of distributed practice: VII. Learning and retention of serial nonsense lists as a function of intralist similarity. *J. exp. Psychol.*, 1952, **44**, 80–87.
5. Underwood, B. J. Studies of distributed practice: XI. An attempt to resolve conflicting facts on retention of serial nonsense lists. *J. exp. Psychol.*, 1953, **45**, 355–359.
6. Underwood, B. J. Studies of distributed practice: XII. Retention following varying degrees of original learning. *J. exp. Psychol.*, 1954, **47**, 294–300.
7. Underwood, B. J., & Richardson, J. Studies of distributed practice: XIII. Interlist interference and the retention of serial nonsense lists. *J. exp. Psychol.*, 1955, **50**, 39–46.
8. Ward, L. B. Reminiscence and rote learning. *Psychol. Monogr.*, 1937, **49**, No. 4 (Whole No. 220).

(Received July 30, 1956)

Journal of Experimental Psychology
Vol. 54, No. 2, 1957

TWO DETERMINANTS OF ASSOCIATIVE REACTION TIME [1]

JERRY S. WIGGINS [2]

Indiana University

This study represents an attempt to investigate, within the framework of the association method, the influences of two factors which are held to be of importance in the formation of verbal associations: (*a*) frequency with which a verbal stimulus has been paired with one or more verbal responses, and (*b*) number of different response alternatives that have been paired with that stimulus.

The association method, in which *S* is instructed to respond verbally to a single verbal stimulus, is considered to be a means of assessing the relative "strength" of associative ties in a hierarchy of competing verbal responses (**8**), with reaction time as the index of response strength. A severe limitation of most association experiments is that the variables which are thought to contribute to the formation of associations (contiguity, frequency, recency, etc.) are not brought under experimental control for a given *S*. Instead, inferences about a given *S*'s history are made from normative data obtained with a group of *S*s (**2, 4, 6, 10**).

Such studies of group associative tendencies have revealed an inverse relationship between response frequency and associative reaction time (**14**). More recently Laffal (**6**) has demonstrated a positive relationship between number of response alternatives and associative reaction time. The nature of the relationship between stimulus frequency and associative reaction time has not been directly determined, however, and evidence from related experiments is contradictory. On the one hand, experiments have indicated that stimulus frequency is positively related to ease of learning (**9**), ease of recall (**4**), and ease of recognition (**11**)—all of which might be expected to be related to decreased latency values in association. On the other hand, other experiments have demonstrated a positive relationship between stimulus frequency and number of response alternatives (**2, 7, 12, 15**) which would suggest increased latency values in association because of the positive relationship existing between number of response alternatives and associative reaction time (**6**). The present study attempts to resolve this apparent contradiction by use of an experimental procedure, similar to Camp's (**1**), in which a miniature verbal repertoire is established in *S*s prior to the association test.

METHOD

Subjects.—The *S*s were 270 male and female students who were unpaid volunteers from introductory psychology classes at Indiana University. They were assigned to one of 27 subgroups in the order in which they appeared.

Apparatus.—The experiment was conducted in a small room soundproofed sufficiently to insure that external sounds would not trip off the sensitive voice-key adjustment. Directly over the experimental table was a 7½-w. bulb which provided the low illumination necessary for a one-way vision mirror. The *S* was seated at a

[1] This study is based on a dissertation submitted to the Psychology Department of Indiana University in partial fulfillment of the requirements for the Ph.D. degree in psychology. The writer is greatly indebted to Dr. A. M. Binder under whose direction the investigation was conducted.

[2] Now at the University of Rochester.

table facing a plain black 47 × 24-in. discrimination panel with a single 6 × 3-in. opening containing a one-way vision mirror. The only other apparatus visible to *S* was a goose-neck microphone bolted to the table directly in front of him.

On *E*'s side of the panel was a stimulus enclosure unit containing a compartment for the stimulus card which could be illuminated by a 7½-w. bulb. A one-way vision mirror was located between the stimulus enclosure and the discrimination panel so that illumination of the stimulus card compartment made the card visible to *S* through the opening in the panel. Next to the stimulus enclosure, and to the right of *E*, were the components of an electronic voice-key arrangement: an audio amplifier, a Standard Electric timer, a holding relay, and a telegraph key. On each trial *E* closed a switch which simultaneously activated a clock and lighted the stimulus enclosure in which *S* saw a stimulus card through the one-way vision mirror. The verbal response of *S* into the microphone stopped the clock. The light was turned off manually by *E*.

Stimulus materials.—Two nonsense syllables were typed side by side on each 8 × 5 in. card with a large-type typewriter and were separated by a small partition on the stimulus screen. Eleven syllables of 40% association value were selected from Hilgard's (**5**) presentation of Glaze's list: VEK, YAN, and BAZ served as stimulus syllables and TIJ, MOQ, GUD, CUG, PEX, FIK, HUY, and KOG were used as response syllables.

Procedure.—The task was structured to *S* as research on the problem of communicating coded information between two people. The *S* was told to look at both of the "code words" on a card and to pronounce the word on the right. Depending upon the treatment group, *S* was then given 8, 16, or 32 such trials. The exposure interval was determined by *S*'s reaction time, which seldom exceeded 4 sec., plus an additional 2 sec. during which the card was exposed after *S* had responded. The intertrial interval was approximately 7 sec.

At the end of the trial series *S* was instructed to respond to a stimulus syllable on the left of the stimulus card with the first other syllable that came to mind. A single test trial was then given and the response and reaction time recorded.

Experimental design.—As shown in Fig. 1, a thrice-replicated two-factor design was employed with frequency of paired presentation and number of response alternatives represented in the row (A) and column (B) treatments, respectively, and the replications to control for the use of idiosyncratic stimulus syllables as the three

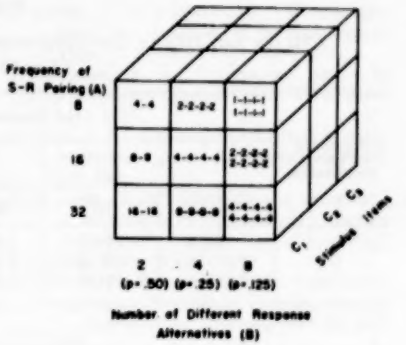

FIG. 1. Experimental design.

layers (C). There were 10 *S*s in each of 27 experimental groups.

The hyphenated numbers on the front of each block refer to the frequency with which a given response alternative appeared in a training series. In Group 4-4, for example, two response alternatives were paired with the stimulus and each alternative was paired four times. Group 2-2-2-2 received four response alternatives and each of the alternatives was paired twice, etc.

Figure 1 also shows the two restrictions that were imposed upon the factorial design to avoid confounding of the two main treatment variables: (*a*) the number of trials (frequency of paired presentation) was held at three different constant frequencies (**8**, **16**, **32**) across columns, and (*b*) the probability of occurrence of a given response alternative (relative frequency of occurrence/number of possible outcomes) was held at three constant values (.50, .25, .125) across rows.

RESULTS AND DISCUSSION

Because of the assumption of normality of distribution in analysis of variance, the reaction time data were converted to log-centisecond form. Table 1 presents the means of the transformed scores for the groups classified by the two main treatment variables. The differences are, with one exception, all in the expected direction. For each level of frequency of paired presentation, reaction time increased as the number of response alternatives increased; for each level

TABLE 1

MEANS AND *SD*'S IN LOG-CENTISECONDS FOR THE MAIN TREATMENT GROUPS ($N = 30$ each group)

Frequency of Paired Presentation	Number of Response Alternatives					
	2		4		8	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
8	2.18439	.11925	2.26911	.15395	2.35221	.20174
16	2.15933	.18614	2.18365	.11979	2.28364	.19644
32	2.11684	.14450	2.19951	.15100	2.16693	.17706

of number of response alternatives, reaction time decreased as frequency of paired presentation increased. The single exception to these trends is found in the 32-trial group which had four response alternatives. This group had somewhat longer reaction times than would be expected from the trends evident in adjoining groups.

An analysis of variance was performed on the transformed scores using the computational procedures described by Edwards (**3**) for complex factorial designs. Table 2 is a summary of this analysis. Frequency of paired presentation and number of response alternatives were significant well beyond the .001 level. The control variable of stimulus items does not approach significance nor do any of the interactions.

TABLE 2

ANALYSIS OF VARIANCE OF ASSOCIATIVE REACTION TIMES (LOG-CENTISECONDS) ON TEST TRIAL

Source of Variation	*df*	*MS*	*F*
Frequency (A)	2	.26137	9.11*
Response alternatives (B)	2	.29530	10.29*
Stimulus items (C)	2	.03489	1.22
A × B	4	.04915	1.71
A × C	4	.02670	0.93
B × C	4	.03866	1.35
A × B × C	8	.01348	0.47
Within groups	243	.02870	
Total	269		

* $P < .001$.

The results of the frequency variable are a confirmation of "Marbe's Law" (**13**) for the individual and may represent a more general principle. That is, the fact that group-response frequencies are negatively related to group-response latencies may be a consequence of the fact that individuals' previous S-R histories with the words involved tend to be reflected in group norms.

The findings with respect to number of response alternatives are consistent with Laffal's (**6**) and provide a somewhat more precise demonstration of the phenomenon in that the entire range of reaction times was studied in relation to associative response hierarchies which were predetermined for each *S*. Taken together, the studies suggest that a relationship of some generality is involved which has implications for both word association tests and the more general problem of the formation of verbal associations.

The consistent trends among means at different levels of the two main treatment variables reflect the absence of interaction effects found in the analysis of variance. This lack of interaction suggests that the facilitating effect upon reaction time of increasing the frequency of paired presentation is independent of the number of response alternatives involved and that the depressing effect upon reaction

time of increasing the number of response alternatives is independent of the particular level of frequency of paired presentation employed.

The extent to which we may extrapolate from these highly limited and simplified verbal hierarchies to the more complex and "meaningful" hierarchies involved in word association is an open question. The consistency of the present findings and the analytic possibilities inherent in the approach suggest that this is a fruitful area for further exploration.

Summary

Two factors which are held to be of importance in the formation of verbal associations were investigated by "building-in" differing associational response hierarchies to a single stimulus. Frequency with which a verbal stimulus is paired with one or more verbal responses and number of different response alternatives that have been paired with that stimulus were independently varied in a factorial design involving 270 college students. Pairs of nonsense syllables were presented visually and responded to orally during a training series. The associative latency of the test trial, in which *S* was instructed to associate to the training stimulus, was the dependent variable.

It was found that: (*a*) associative latency decreases as frequency of S-R pairing increases; (*b*) associative latency increases as the number of response alternatives increases; and (*c*) when frequency of S-R pairing and number of response alternatives are varied independently within a single design, there are no significant interaction effects upon associative latency.

References

1. Camp, B. W. The formation and recall of associations as a function of syllable familiarity and frequency of pairing. Unpublished Ph.D. thesis, Indiana Univer., 1954.
2. Cofer, C. N., & Shevitz, R. Word-association as a function of word-frequency. *Amer. J. Psychol.*, 1952, **65**, 75–79.
3. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1951.
4. Hall, J. F. Learning as a function of word-frequency. *Amer. J. Psychol.*, 1954, **67**, 138–140.
5. Hilgard, E. R. Methods and procedures in the study of learning. In S. S. Stevens (Ed.) *Handbook of experimental psychology.* New York: Wiley, 1951.
6. Laffal, J. Response faults in word association as a function of response entropy. *J. abnorm. & soc. Psychol.*, 1955, **50**, 265–270.
7. Lepley, W. M. An hypothesis concerning the generation and use of synonyms. *J. exp. Psychol.*, 1950, **40**, 527–530.
8. Noble, C. E. An analysis of meaning. *Psychol. Rev.*, 1952, **59**, 421–430.
9. Noble, C. E. The effect of familiarization upon serial verbal learning. *J. exp. Psychol.*, 1955, **49**, 333–338.
10. Schlosberg, H., & Heineman, C. The relationship between two measures of response strength. *J. exp. Psychol.*, 1950, **40**, 235–247.
11. Solomon, R. L., & Postman, L. Frequency of usage as a determinant of recognition thresholds for words. *J. exp. Psychol.*, 1952, **43**, 195–201.
12. Thorndike, E. L., & Lorge, I. *The teacher's word book of 30,000 words.* New York: Bur. Pub., Teacher's Coll., Columbia Univer., 1944.
13. Thumb, A., & Marbe, K. *Experimentelle Untersuchungen uber die psychologischen Grundlagen der sprachlichen Analogiebildung.* Leipzig: Englemann, 1901.
14. Woodworth, R. S., & Schlosberg, H. *Experimental psychology.* Rev. ed. New York: Holt, 1954.
15. Zipf, G. K. The meaning-frequency relationship of words. *J. gen. Psychol.*, 1945, **33**, 251–256.

(Received July 30, 1956)

Journal of Experimental Psychology
Vol. 54, No. 2, 1957

THE EFFECT OF AN IRRELEVANT DRIVE ON MAZE LEARNING IN THE RAT

HARRY W. BRAUN, CARL E. WEDEKIND, AND JOSEPH F. SMUDSKI

University of Pittsburgh and Montefiore Institute of Research

Hull (**4**) has made the assumption that response strength is a multiplicative function of learning or cognitive factors and motivational or drive factors. Drive factors are designated by the hypothetical construct, drive strength (D), which comprises the total effective drive strength operating in S at any given moment. In any experimental situation, both relevant and irrelevant needs or motivation are assumed to summate to contribute to the value of D. A relevant need is one which is reduced by the response under consideration; irrelevant needs encompass all other primary and secondary needs present in S when the response occurs but which are not reduced by the response. $\bar{D}$ and $\hat{D}$ denote drive strengths ascribable to relevant and irrelevant need states respectively, and the manner in which they contribute to D has been tentatively defined by Hull as follows: $D = (\bar{D} + \hat{D})/(\hat{D} + 100)$.

These considerations have stimulated research with the rat to determine the conditions under which and the manner in which relevant and irrelevant drives combine to affect behavior. In general, it has been found that a relevant and an irrelevant drive summate to yield a greater drive strength than will either one alone.

Thus, Kendler (**5**) investigated the acquisition of a lever-pressing response for food under combinations of irrelevant drive values resulting from 3, 6, 12, and 22 hr. of water deprivation with a constant relevant drive value resulting from 22 hr. of food deprivation. Relevant and irrelevant drive values were found to summate to increase total drive state up to and including the irrelevant drive value of 12 hr. The combination of 22-hr. water deprivation and 22-hr. food deprivation resulted in a response strength that was less than that obtained when only food deprivation was the motivational condition.

Siegel (**7**) conditioned a lever-pressing response for food under a 22-hr. hunger drive and then extinguished this response under a drive condition of 22-hr. food deprivation combined with 22 hr. water deprivation for one subgroup and under only a 22-hr. hunger drive for the other subgroup. Although the latter subgroup required more trials to reach the extinction criterion, the difference between the two subgroups was not statistically significant.

Webb (**10**) trained rats under hunger drive to push against a door to obtain food, and after the completion of training he satiated them for food. *Ss* were then divided into four subgroups which were deprived of water for 0, 3, 12, or 22 hr. In subsequent trials to extinction, response strength was found to be directly related to the strength of the irrelevant need.

Amsel (**1**) combined an irrelevant hunger need resulting from 22 hr. of food deprivation with, in one case, a relevant drive to escape pain from electric shock stimulation (two levels), and in a second case, a relevant secondary need to escape anxiety (two levels) which was defined as a conditioned pain reaction. The former combination had no apparent effect on speed of running, while the latter significantly increased response speed.

Common to these four studies using rat *Ss* is a relatively simple response, and the general trend of the results is that two needs may summate to increase response strength. It will also be noted that in these investigations drive summation has been studied in terms of two appetitional needs (hunger and thirst) either in combination with one another or with pain and anxiety. The need for additional research involving complex tasks, other drives, and other drive values is apparent.

The present experiment was de-

signed to provide information on the effect of drive combination on the acquisition by the rat of a complex task: a maze habit. Maze learning was measured as a function of a combination of a relevant drive—escape from water—and an irrelevant drive—hunger resulting from food deprivation. A water temperature of 15 ± 1° C. was chosen for the high level of motivation and a temperature of 35 ± 1° C. for the low level of motivation. The irrelevant hunger drive was induced by 0- and 22-hr. food deprivation. Wever (**11**) has shown that rats swim more rapidly when the temperature of the water is 10° to 15° C. than when the water is near body temperature (37.50° C). Similarly, Hack (**3**) and Pilgrim, Zabarenko, and Patton (**6**) have demonstrated the elimination of errors in maze learning is more rapid at the former temperatures than at the latter.

Method

Learning problem.—A modified, 5-choice-point Lashley III maze adapted for swimming was used (**6**). Doors were operated to prevent retracing. On each trial *S* was lowered into the maze at the start position and then swam to the mesh-screen landing platform which was above the water level. Total swimming time and errors were recorded on each trial. An error was defined as entrance into a blind alley or retreat from a choice point by a full body length, exclusive of the tail. All *Ss* were given 15 maze trials with an intertrial interval of 24 hr. Prior to the first maze trial, each *S* was placed in the water and was required to swim the length of one alley to the landing platform: once the first day, twice the second day, and so forth, until a total of 15 preliminary trials had been given through Day 5.

Subjects.—The *Ss* were 62 male rats of the Sprague-Dawley strains, ranging in age from 108 to 123 days at the beginning of the experiment. The *Ss* were randomly divided into groups representing the four experimental conditions: (*a*) 35° C.-22 hr. of food deprivation ($N = 15$); (*b*) 35° C.-0 hr. of food deprivation ($N = 15$); (*c*) 15° C.-22 hr. of food deprivation ($N = 16$); and (*d*) 15° C.-0 hr. of food deprivation ($N = 16$).

The 22-hr. irrelevant drive groups were fed when a minimum of 30 min. had elapsed after the last *S* of the group had completed his daily maze trial. They were then allowed to feed for 1 hr. The *Ss* in the other two groups had access to food at all times except during maze trials. Water was always available in the living cages of all *Ss*.

Results

The results of a repeated measurements analysis of variance of log swimming times are presented in Table 1. The differences among the groups are significant beyond the .01

TABLE 1

Analyses of Variance of Log Swimming Times and Error Scores

Source of Variation	*df*	Time		Errors	
		MS	*F*	*MS*	*F*
Between groups	3	7.769	10.08**	316.29	9.75**
Temperature (Te)	1	15.829	20.53**	797.29	24.58**
Deprivation (De)	1	5.662	7.34**	100.03	3.08
Te × De	1	1.817	2.36	51.54	1.59
Between *Ss* within groups	58	0.771		32.43	
Trials (Tr)	14	2.290	41.64**	773.53	95.03**
Groups × Tr	42	0.241	4.38**	15.32	1.88*
Tr × Te	14	0.502	9.13**	15.07	1.85*
Tr × De	14	0.087	1.58	9.50	1.17
Tr × Te × De	14	0.134	2.44**	21.39	2.63**
Between *Ss* within Groups × Tr	812	0.055		8.14	

* $.05 > P > .01$.
** $P < .01$.

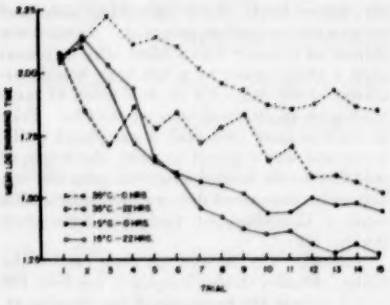

FIG. 1. Mean log swimming time on the 15 maze trials.

level. The extent and direction of group differences may be seen in Fig. 1 which contains the mean log swimming times of the four groups on the 15 maze trials. The application of *t* tests showed that the 15° C.-22 hr. group swam significantly faster than the 15° C.-0 hr. group on Trials 8, 9, 12, 13, 14, and 15. Similar tests revealed the significant superiority of the 35° C.-22 hr. group over the 35° C.-0 hr. group on all maze trials except 1, 9, and 11. On Maze Trials 2, 3, 6 through 11, and 13, the 15° C.-0 hr. group was significantly faster than the 35° C.-22 hr. group. The 15° C.-22 hr. group swam significantly faster than the two 35° C. groups on at least 12 of the trials, and the 15° C.-0 hr. group was superior to the 35° C.-0 hr. group on 13 of the trials.

Table 1 also shows that the main effects of water temperature and food deprivation are significant beyond the .01 level. The interaction of these variables, however, is not significant, which indicates that as far as total learning is concerned, the relevant and irrelevant drives combine additively rather than multiplicatively. As expected, differences between trials is significant at the .01 level. The interaction, Groups × Trials is significant at the .01 level and when this term is broken into its three components it is found that the Trials × Temperature interaction and the triple interaction of Trials × Temperature × Deprivation are both significant at the .01 level. The Trials × Deprivation interaction is not significant at an acceptable level. To account for the significance of the triple interaction, the significance of the difference in log swimming times on each maze trial between the two temperature groups under 0-hr. food deprivation and the two temperature groups under 22-hr. deprivation was tested. The resulting *t*'s were significant beyond the .01 level on Trials 2 through 6. None of the other *t*'s was significant at the .05 level.

The results of a repeated measurements analysis of variance of errors also appear in Table 1. The pattern of significance among the main effects and their various interactions is much the same as in the case of log swimming times, but errors seem to be a less sensitive measure of behavior in this experiment. The mean error scores of the four groups on the 15 maze trials are presented in Fig. 2.

An analysis of the difficulty of the maze appears in Table 2 which shows the number of *S*s in each of the groups who made the correct response at each of the choice points on the first maze trial. The greater strength of the incorrect response at each choice point, as measured in this manner, is apparent. Differences among the groups in terms of this measure are not significant when Cochran's *Q* test was applied (**8**, pp. 161–165), indicating the initial equality of incorrect response tendency among the groups.

At the start of the experiment, differences in weight among the four groups were not significant, but at the conclusion of the 15 maze trials the 0-hr. deprivation groups were sig-

nificantly heavier than the other two groups.

Discussion

The major finding of this experiment is that the addition of an irrelevant drive based on 22-hr. food deprivation to both high and low levels of a relevant drive to escape from water resulted in an increased total drive strength as measured by swimming time in a 5-choice-point maze. An inspection of Fig. 1 and 2 also suggests that at the low level of relevant motivation (35° C.) the effect of the addition of the irrelevant drive was greater than at the high level of relevant motivation (15° C.) This observation is consistent with the implication of Hull's formulation of drive summation cited above. From the data presented in this study, it also appears that differences in learning attributable to the addition of the irrelevant drive are greater near the end of the series of 15 maze trials than at the beginning. The results also indicate that at a high level of relevant drive without the addition of an irrelevant drive (15° C.-0 hr. group), learning is superior to that associated with a low relevant drive level combined with a high irrelevant drive level (35° C.-22 hr. group).

Against the background of the studies by Kendler (**5**), Webb (**10**), Siegel (**7**), and Amsel (**1**), wherein the effects of drive summation on the acquisition or extinction of a single response were investigated, this experiment represents an extension which shows that drive summation may facilitate multiple-response learning. On the face of it, this finding is not in agreement with the data of human trial-and-error (multiple response) learning (**2, 9**) which show that a heightened drive state producedthrough a combination of relevant and irrelevant drives (manifest anxiety) hinders learning. However, the authors recognize that the experimental variables, including the maze task, employed in this investigation probably exceed the boundary conditions of the drive theory which is being formulated in this latter group of studies.

TABLE 2

Number of *Ss* Who Made the Correct Response at the Five Choice Points on Trial 1

Group	*N*	Choice Point					Total
		I	II	III	IV	V	
35° C.-0 hr.	15	2	1	1	1	2	7
35° C.-22 hr.	15	0	1	1	1	2	5
15° C.-0 hr.	16	2	2	1	1	3	9
15° C.-22 hr.	16	1	2	1	2	1	7
Total	62	5	6	4	5	8	28

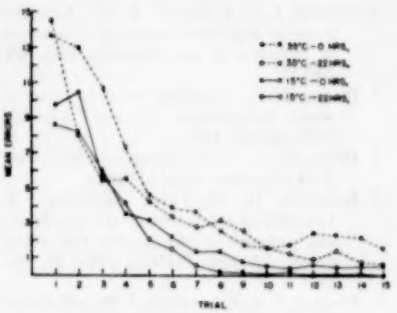

Fig. 2. Mean errors on the 15 maze trials.

Summary

This study investigated the learning by the rat of a modified Lashley Maze III adapted for swimming as a function of the combination of a relevant drive—escape from water— and an irrelevant drive—hunger resulting from food deprivation. Two levels of relevant drive were produced by using water temperatures of 15° C. (high) and 35° C. (low). The irrelevant drive was induced by 0-hr. and 22-hr. food deprivation. The addition of a high irrelevant drive to both high and low levels of relevant drive resulted in an increased total drive strength as measured by swimming time and error elimination.

References

1. Amsel, A. The combination of a primary appetitional need with primary and secondary emotionally derived needs. *J. exp. Psychol.*, 1950, **40**, 1–14.

2. Farber, I. E., & Spence, K. W. Complex learning and conditioning as a function of anxiety. *J. exp. Psychol.*, 1953, **45**, 120–125.
3. Hack, E. R. Learning as a function of water temperature. *J. exp. Psychol.*, 1933, **16**, 442–445.
4. Hull,, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.
5. Kendler, H. H. Drive interaction: I. Learning as a function of the simultaneous presence of the hunger and thirst drives. *J. exp. Psychol.*, 1945, **35**, 96–109.
6. Pilgrim, F. J., Zabarenko, L. M., & Patton, R. A. The role of amino acid supplementation and dietary protein level in serial learning performance of rats. *J. comp. physiol. Psychol.*, 1951, **44**, 26–36.
7. Siegel, P. S. Alien drive, habit strength, and resistance to extinction. *J. comp. Psychol.*, 1946, **39**, 307–317.
8. Siegel, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.
9. Taylor, J. A., & Spence, K. W. The relationship of anxiety level to performance in serial learning. *J. exp. Psychol.*, 1952, **44**, 61–64.
10. Webb, W. B. The motivational aspect of an irrelevant drive in the behavior of the white rat. *J. exp. Psychol.*, 1949, **39**, 1–14.
11. Wever, E. G. Water temperature as an incentive to swimming activity in the rat. *J. comp. Psychol.*, 1932, **14**, 219–224.

(Received July 30, 1956)